

# **A review of the Longitudinal Study of Young People in England (LSYPE): recommendations for a second cohort**

**Aleks Collingwood, Hayley Cheshire, Gerry Nicolaas, John D'Souza, Andy Ross, Julia Hall, Cole Armstrong, Amanda Prosser, Rosie Green, Debbie Collins, Michelle Gray, Carol McNaughton Nicholls**

This research report was commissioned before the new UK Government took office on 11 May 2010. As a result the content may not reflect current Government policy and may make reference to the Department for Children, Schools and Families (DCSF) which has now been replaced by the Department for Education (DFE).

The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education.

# Contents

<b>List of abbreviations</b> .....	<b>1</b>
<b>Acknowledgements</b> .....	<b>3</b>
<b>Executive summary</b> .....	<b>4</b>
Introduction .....	4
Methodology .....	5
Key recommendations for each objective of the review .....	5
<b>Introduction to the review</b> .....	<b>9</b>
Methodology .....	10
Literature reviews .....	11
Analysis of first LSYPE cohort data .....	11
Telephone interviews .....	11
Workshops with young people .....	12
Depth interviews .....	13
User Group Seminar .....	13
<b>2 Overview of LSYPE and rationale for a second cohort</b> .....	<b>15</b>
2.1 Overview of LSYPE .....	15
Objectives of LSYPE .....	15
Who runs LSYPE? .....	15
How long has the study been running and what stage is it at? .....	16
How has the study altered between Waves? .....	16
What does the LSYPE questionnaire cover? .....	16
2.2 How has LSYPE been used both within DfE and externally? .....	17
How have the data been used? .....	18
Published analyses .....	18
External analysis being undertaken .....	19
Other uses of LSYPE data .....	19
2.3 The benefits of a second cohort and evidence gaps .....	20
Content of the first LSYPE cohort and how this compares to other studies .....	20
2.4 Value for money .....	23
Additional topics that a second cohort could cover .....	26
Improvements on existing information collected .....	28
2.5 Analysis of a second cohort .....	29
2.6 Advisory Panels .....	34
The Previous LSYPE Advisory Panels .....	34
Recruiting stakeholders into an Expert Advisory Panel .....	35
Recruiting LSYPE cohort members into a Young Person Advisory Panel .....	35
Areas for Consultation .....	37
2.7 Recommendations .....	37
<b>3 Sample and fieldwork design</b> .....	<b>40</b>
3.1 Strengths and weaknesses of LSYPE sample .....	40
3.2 The target and study populations and exclusions .....	41
3.3 Sampling frames and clustering .....	42
Maintained Schools .....	42
Independent schools .....	44
3.4 Boosts and sample sizes .....	44
3.5 Stratification .....	46
3.6 Age of cohort and fieldwork strategy .....	46
Fieldwork timing .....	48
3.7 Recommendations .....	49
<b>4 Response rates</b> .....	<b>51</b>
4.1 Explaining response rates .....	51
4.2 LSYPE study response rates .....	52
Cross-sectional response rates .....	52
Longitudinal response rates .....	53
4.3 A comparison of respondents and non-respondents .....	55

4.4 Predictors of non-response .....	56
4.5 Motivation for participation or refusal – views from respondents .....	59
4.6 Recommendations.....	62
<b>5 Mixed modes .....</b>	<b>64</b>
5.1 Strengths and weaknesses of the first LSYPE cohort mixed mode design.....	64
5.2 What are the alternative mixed mode designs? .....	64
5.3 Characteristics of different modes.....	66
5.4 Respondents' mode preference .....	69
5.5 Mixed mode approaches and the LSYPE.....	71
5.6 Recommendations.....	73
<b>6 The use of incentives .....</b>	<b>74</b>
6.1 Overview of key longitudinal studies and their incentive methods .....	74
6.2 The effectiveness of different incentive strategies.....	76
6.3 Recommendations.....	78
<b>7 Interviewing parents.....</b>	<b>80</b>
7.1 Who should be interviewed? .....	80
Who was interviewed at each Wave of the first LSYPE cohort.....	80
7.2 Other longitudinal studies that include a parental interview .....	81
Data collection methods with parents .....	81
What are the parent interviews used for? .....	82
7.3 Including a second parent interview .....	82
7.4 Recommendations.....	83
<b>8 Data linkage.....</b>	<b>85</b>
8.1 Strengths and weaknesses of first LSYPE cohort and data linkage .....	85
8.2 Benefits of data linkage .....	86
8.3 Potential sources of data linkage for LSYPE.....	87
8.4 Respondent views on data linkage.....	92
8.5 Recommendations.....	93
<b>9 Funding.....</b>	<b>95</b>
9.1 Funding of LSYPE .....	95
9.2 Cross-departmental funding and lessons learned.....	95
9.3 Other cross-Government funded studies.....	96
9.4 Disadvantages of cross-departmental funding .....	97
9.5 Recommendations.....	97
<b>10 Communication plan .....</b>	<b>99</b>
10.1 Communicating with respondents .....	100
10.2 Distributing findings and engaging users .....	103
10.3 Communication Plan - Recommendations.....	105
<b>11 Conclusion .....</b>	<b>107</b>
11.1 Objectives of the review and how these have been fulfilled.....	107
11.2 Essential and desirable elements of a new LSYPE cohort .....	109
11.3 For further discussion.....	114
11.4 Final points.....	114
<b>References.....</b>	<b>116</b>
<b>Appendix A Advisory Panels of the first LSYPE cohort.....</b>	<b>121</b>
<b>Appendix B List of content of LSYPE (questionnaire) .....</b>	<b>124</b>
<b>Appendix C Publications on LSYPE.....</b>	<b>129</b>
<b>Appendix D Content and review of other datasets .....</b>	<b>135</b>
<b>Appendix E Youth policies that may affect the analysis questions .....</b>	<b>142</b>
<b>Appendix F Cohort studies with Respondent Advisory Panels.....</b>	<b>144</b>
<b>Appendix G Response Rates .....</b>	<b>146</b>

<b>Appendix H</b>	<b>Details of fieldwork years for other studies .....</b>	<b>179</b>
<b>Appendix I</b>	<b>Expert Panel review of content of LSYPE.....</b>	<b>180</b>
<b>Appendix J</b>	<b>Theories of exchange .....</b>	<b>187</b>
<b>Appendix K</b>	<b>Monetary Incentives.....</b>	<b>189</b>
<b>Appendix L</b>	<b>Parental Interviews.....</b>	<b>193</b>
<b>Appendix M</b>	<b>Communication plan.....</b>	<b>198</b>

## **Tables**

Table 0.1	Summary of recommendations for a second LSYPE cohort.....	6
Table 0.2	Summary of recommendations for a second LSYPE cohort continued ...	7
Table 0.3	Summary of recommendations for a second LSYPE cohort continued ...	8
Table 1.1	Summary of characteristics of young people who were interviewed.....	12
Table 2.1	Future topics for a second cohort as recommended by stakeholders and users.....	27
Table 3.1	Advantages of using schools or geographical areas as the clustering variable.....	43
Table 4.1	Cross-sectional response rates .....	52
Table 4.2	Longitudinal response rates.....	53
Table 5.1	Summary of the strengths and weaknesses of different modes of data collection .....	67
Table 8.1	Sources of information used in a selection of English longitudinal studies.....	86
Table 11.1	Recommendations for a second LSYPE cohort .....	110
Table 11.2	Recommendations for a second LSYPE cohort .....	112
Table 11.3	Recommendations for a second LSYPE cohort .....	113
Table 11.4	Areas for further discussion .....	114

# List of abbreviations

ALSPAC	Avon Longitudinal Study of Parents and Children
BCS70	1970 British Cohort Study
BES	British Election Study
BHPS	British Household Panel Study
BIS	Department for Business, Innovation and Skills (2009-present, previously known as DIUS)
BME	British Minority Ethnic
CAPI	Computer Assisted Personal Interviewing (face to face)
CATI	Computer Assisted Telephone Interviewing
CAWI	Computer Aided Web Interviewing
CLG	Department for Communities and Local Government
CWI	Child Wellbeing Index
DCSF	Department for Children, Schools and Families (2007-2010, now known as DfE)
DEFRA	Department for Environment, Food, and Rural Affairs
DfE	Department for Education (2010-present, previously known as DCSF)
DIUS	Department for Innovation, Universities and Skills (2007-2009)
DWP	Department for Work and Pensions
ELSA	English Longitudinal Study of Ageing
EPPSE	Effective Pre-School, Primary and Secondary Education (EPPSE 3-14)
ESDS	Economic and Social Data Service
ESRC	Economic and Social Research Council
FACS	Families and Children Study
FSM	Free School Meals
GCSE	General Certificate of Secondary Education
GHQ12	The 12-item General Health Questionnaire
GOR	Government Office Region
GUS	Growing Up in Scotland study
HES	Hospital Episode Statistics
HESA	Higher Education Statistics Agency
HILDA	Household Income and Labour Dynamics in Australia survey
HMRC	HM Revenue and Customs
HRS	Health and Retirement Study
IDACI	Income Deprivation Affecting Children Index
IFS	Institute of Fiscal Studies
IMD	Index of Multiple Deprivation
IRAS	Integrated Research Application System
LSAC	Longitudinal Study of Australian Children (Growing Up in Australia)
LSYPE	Longitudinal Study of Young People in England
KS3	Key Stage 3
MCS	Millennium Cohort Study

MoJ	Ministry of Justice
MRIS	Medical Research Information Service
NatCen	National Centre for Social Research
NCVO	National Council for Voluntary Organisations
NCDS	National Child Development Study
NEET	Not in Education, Employment or Training
NIA	National Institute on Aging (US research centre)
NIGB ECC	National Information Governance Board for Health and Social Care's Ethics and Confidentiality Committee
NHS	National Health Service
NHS IC	National Health Service Information Centre
NLSY	National Longitudinal Surveys of Youth
NLSCY	National Longitudinal Survey of Children and Youth
NPDDU	National Pupil Database and Dissemination Unit
NPIA	National Policing Improvement Agency
NPD	National Pupil Database
OMSAS	Offender Management and Sentencing Analysis Services
ONS	Office for National Statistics
PIAP	PNC Access Panel
PLASC	Pupil-level Annual Schools Census
PLUG	PLASC/ NPD Users' Group
PNC	Police National Computer
PSID	Panel Study of Income Dynamics
PSU	Primary Sampling Unit
RPA	Raising the Participation Age
REC	Research Ethics Committee
SHA	Strategic Health Authority
SIPP	Survey of Income and Program Participation
SLID	Canadian Survey of Labour and Income Dynamics
SHP	Swiss Household Panel Study
US	Understanding Society
YCS	Youth Cohort Study

## Acknowledgements

This research report has many different elements to it and has consequently had a large number of people involved in different areas of the research. I would like to thank everyone who took part in the review, including all authors, PublicZone, current cohort members, young people who took part in workshops and all stakeholders who attended the User Group Seminar. In particular I would like to thank Hayley Cheshire from the National Centre for Social Research (NatCen) for her commitment, organisation and guidance as the original Project Manager for this review. At the Department for Education (DfE), I would like to thank Clare Baker and Helen Wood for their guidance on the research and Phil Emmott for his input into the final report. I would also like to thank Carli Lessof, Bob Erens and Elizabeth Hacker from NatCen, Martha Paren from PublicZone, Nick Moon from GFK-NOP, Clare Crawford from IFS, and Nicholas Howat, Mark Peters and Carrie Harding from TNS-BMRB.

Aleks Collingwood  
LSYPE Review Project Manager  
National Centre for Social Research

The Department for Education (DfE) replaced the Department for Children Schools and Families (DCSF) in May 2010. To ensure consistency DfE is referred to throughout this report regardless of whether this is being referred to retrospectively or currently. Previous names (the Department for Children, Schools and Families, the Department for Education and Skills and the Department for Education and Employment) are retained in references to printed material issued before the inception of the DfE, and in other contexts where it makes sense.

# Executive summary

## Introduction

The UK is internationally renowned for its longitudinal studies and the high quality data they generate to inform and assess policy. The Longitudinal Study of Young People in England (LSYPE) is a unique and highly successful example of one of these studies, which generates large volumes of policy-relevant data for analysis. LSYPE follows the transitions of a representative cohort of young people in England into adulthood in greater depth than any other existing data source.

Since the launch of the first cohort of the LSYPE in 2004 there have been significant changes to the economic and political environment in which young people make their transitions. Significantly, in the year the seventh Wave of the first LSYPE cohort concludes, there has also been a change of government (in May 2010). There is therefore a clear need to study a new (second) cohort of young people who are being affected by these changes, and who will reach adulthood in this very different economic and political environment. This second cohort could also be compared with the first cohort, an additional dimension – both in terms of analytical value and also return on investment - to the current LSYPE cohort. This review aims to support the case for a second LSYPE cohort by reviewing in detail the processes in place during the first LSYPE cohort and making recommendations for improvements to these so that should a second cohort take place, it is best situated to play a crucial role in informing policy and meeting a wide range of needs.

This review has been published within a few months of the change of Government. At the time of writing the new Government's priorities are being developed, and although it is unclear exactly what form many policies will take, the Government have outlined a number of key policy areas of relevance to children and young people. These areas include encouraging shared parenting, providing greater support to disabled children and those with special educational needs, and protecting children in the event of family breakdown. A second LSYPE cohort would be instrumental in providing a sound evidence base for, and evaluating the success of, these new policies. This report presents recommendations that could form the foundations for planning a second cohort, although any change in policy priorities may impact on the exact content and analysis of this.

## Methodology

The review draws on the expertise of a number of specialists, including academics, senior analysts, survey methodologists, statisticians, communications and policy specialists, as well as young people themselves. A variety of methods were used to undertake this review, including desk research, quantitative analysis, depth interviews, telephone interviews and focus groups. Discussions were held with key staff members involved in LSYPE from the Department for Education and the consortium managing the first cohort, as well as with academics that have used LSYPE data. In order to inform the communication plan for a second cohort, a consultation among first cohort members and other young people was conducted. Finally, a User Group Seminar was held, allowing the opportunity to consult a variety of users before making formal recommendations about sample sizes, mode and other elements of survey design (e.g. face to face interviewing, self completion, telephone interviewing, web interviewing) and to discuss key evidence gaps that a second LSYPE cohort could fill.

## Key recommendations for each objective of the review

In Table 0.1 (below) the key recommendations made pertaining to the design of second LSYPE cohort are summarised. The justification for these recommendations can be found throughout the report:

**Table 0.1 Summary of recommendations for a second LSYPE cohort**

<b>Topic</b>	<b>Recommendations</b>
<b>Content</b> (See page 26)	<ul style="list-style-type: none"> <li>• Retain as much of the content of the first cohort as much as possible to enable change to be monitored between the two cohorts.</li> <li>• Improve the content by including a stronger focus on health, psychological and personality measurements, school choice, career aspirations, parenting, social networks (friends, siblings, contact with absent parents etc) and sexual orientation.</li> </ul>
<b>Analysis plan</b> (See page 29)	<ul style="list-style-type: none"> <li>• Focus on the particular strengths of the dataset, drawing out what makes the LSYPE unique from other sources of data, such as the National Pupil Database.</li> <li>• One of the most obvious strengths would be the possibilities that a second cohort offers for comparative studies with the first LSYPE cohort. Use this to explore change in the experience and context of young people's lives and enable an evaluation of Government policies.</li> <li>• Use longitudinal analysis to explore young people's development over time and to gain a better understanding of the different pathways they follow.</li> <li>• Explore thematic policy strands such as: young people's wellbeing; young people's attainment; young people and school choice; improving the achievement of disadvantaged young people; parental engagement; transitions into further education, training or employment; vocational training; vulnerable young people, young people's engagement in risky activities and the effect of the Big Society (school, parents, community) on supporting young people to make successful transitions.</li> <li>• Policy strands could also be developed around emerging new policy priorities.</li> </ul>
<b>Advisory Panels</b> (See page 34)	<ul style="list-style-type: none"> <li>• Two separate Advisory Panels should be set up to support the development of a second LSYPE cohort: <ul style="list-style-type: none"> <li>○ An Expert Advisory Panel including academics and policy makers</li> <li>○ A Young Person Advisory Panel.</li> </ul> </li> <li>• The Advisory Panels would be consulted separately at the setting up stage of a second cohort and then at timely intervals (such as annually) while the second cohort becomes established.</li> <li>• Advisory Panel members should be consulted on issues such as questionnaire content, use of incentives, methods of keeping in touch with respondents, the design of the study and the best ways of presenting study findings.</li> </ul>
<b>Sample and fieldwork design</b> (See page 40)	<p><b>Maintained schools</b></p> <ul style="list-style-type: none"> <li>• A two stage sample of pupils in maintained schools, through the Annual Schools Census should take place.</li> <li>• Schools should be used as the primary sampling unit (PSU).</li> <li>• As with the first LSYPE cohort, approximately 30 pupils per PSU should be selected.</li> <li>• PSUs should be stratified before sampling. Recommended stratifiers include: type of school, different GOR, urban/rural and academic performance.</li> <li>• Pupils should be stratified by variables. Recommended stratifiers include: ethnicity, gender, FSM, IDACI and KS3 achievement.</li> <li>• Pupils from certain subgroups of interest, such as minority ethnic groups, disadvantaged groups and under achievers, should have an increased selection probability.</li> <li>• Sample sizes should not be chosen to achieve a large Wave 1 sample, but be chosen to achieve an adequate effective sample size for each important subgroup in later Waves.</li> </ul>

**Table 0.2 Summary of recommendations for a second LSYPE cohort continued ...**

<b>Topic</b>	<b>Recommendations</b>
<b>Sample and fieldwork design</b> (See page 40)	<p><b>Independent schools</b></p> <ul style="list-style-type: none"> <li>• A two stage sample of independent pupils using the school as the primary sampling unit (PSU) is recommended.</li> <li>• Following the first LSYPE cohort, approximately 30 pupils per PSU should be chosen.</li> <li>• PSUs should be stratified before sampling. Recommended stratifiers include: whether they are single or mixed sex, GOR and academic performance.</li> <li>• Independent school stratification at pupil level is not recommended.</li> </ul> <p><b>Recommended age of cohort and fieldwork year</b></p> <ul style="list-style-type: none"> <li>• The age of cohort in a second LSYPE should be Year 9 pupils. This is to promote comparability with the first cohort.</li> <li>• Fieldwork should not commence before 2012. This is in order to survey one of the first cohorts likely to experience the full impact of remaining in education until aged 18 (following the implementation of the Education and Skills Act 2008).</li> <li>• The timing of interviews should remain the same as they were for the first LSYPE cohort, with interviews taking place over the spring and summer months. However care should be taken that fieldwork does not clash with exam periods.</li> </ul>
<b>Response rates</b> (See page 51)	<ul style="list-style-type: none"> <li>• Reissuing non-responders in subsequent Waves could improve response rates in a second LSYPE cohort.</li> <li>• Hard to reach groups should be specifically targeted, such as those with lower educated parents and with lower socio-economic status, those from lone parent households and those living in disadvantaged circumstances.</li> </ul>
<b>Mixed modes</b> (See page 64)	<ul style="list-style-type: none"> <li>• The first three Waves should be conducted using face to face interviews with young people.</li> <li>• A sequential mixed mode approach is recommended after Wave 3 with the cheapest mode used first and then more expensive modes worked through for non-responders.</li> <li>• The survey questionnaires should be designed from the outset to be mixed mode, and be designed so that questions with the greatest risk of measurement error between modes are not asked in mixed mode Waves if possible.</li> <li>• Internet interviewing should not be a sole option, as some respondents will not have access or literacy.</li> </ul>
<b>Use of incentives</b> (See page 74)	<ul style="list-style-type: none"> <li>• Unconditional monetary Incentives of approximately £10 should be offered to increase participation in each Wave.</li> <li>• A marginal increase to £15 at Wave 4 is recommended. There is evidence to suggest this marginally improves response rates.</li> </ul>

**Table 0.3 Summary of recommendations for a second LSYPE cohort continued ...**

Topic	Recommendations
<b>Parent interview</b> (See page 80)	<ul style="list-style-type: none"> <li>• Interviews with both the main parents are recommended for the first three Waves and interviews with second parents for at least Wave 1.</li> <li>• Interviews with second parents are also recommended for the first three Waves depending on the desired content of the study and if the resources are available.</li> <li>• Second parent interviews at some or all Waves could be conducted using web-based or postal questionnaires to reduce costs (see mixed modes also).</li> <li>• Testing the accuracy of collecting information by proxy for second parents could also lead to reduced costs.</li> </ul>
<b>Data linkage</b> (See page 85)	<ul style="list-style-type: none"> <li>• It is recommended that data linkage with the National Pupil Database continues.</li> <li>• Data linkage should certainly be pursued with Health Episode Statistics (NHS), the Work and Pensions Longitudinal Study (Department for Work and Pensions) and higher education data (HESA).</li> <li>• Other sources of administrative data could be considered, including criminal records held on the Police National Computer (Ministry of Justice).</li> </ul>
<b>Funding</b> (See page 95)	<ul style="list-style-type: none"> <li>• Careful consideration should be given to the pitfalls of cross-departmental funding in regards to broadening the scope of the study.</li> <li>• If cross-departmental funding is pursued, the distinctiveness and unique benefits of LSYPE need to be clearly defined and communicated to other Departments at the outset.</li> <li>• Cross-sectional analysis of the first wave of the study should be encouraged, which would help to engage potential funding Departments early on in the life of the study.</li> <li>• Ensure that the contributing Departments have opportunities to feed into the questionnaire design.</li> </ul>
<b>Communication plan</b> (See page 99)	<p data-bbox="456 1199 915 1226"><b>Communicating with users and respondents</b></p> <ul style="list-style-type: none"> <li>• If funding allows, appoint a communications manager to offer strategic advice about communications to young people and parents, develop materials, and ensure key messages are communicated consistently</li> <li>• Traffic to study websites and responses to emails and SMS messages should be regularly monitored.</li> <li>• Track who is accessing findings, in what format and how often in order to maximise interaction and improve communication.</li> </ul> <p data-bbox="456 1535 1084 1562"><b>Communicating with respondents and those that effect them</b></p> <ul style="list-style-type: none"> <li>• Prepare a communications plan including target audiences, branding and key messages, materials to be produced, timelines for distribution and measures for success.</li> <li>• Develop and test materials for parents and young people.</li> <li>• Ensure an ongoing communication takes place to keep audiences engaged between Waves.</li> </ul> <p data-bbox="456 1759 776 1787"><b>Communicating to study users</b></p> <ul style="list-style-type: none"> <li>• Prepare a database of report/data recipients.</li> <li>• Develop a single, professional looking and usable website.</li> <li>• Prepare a distribution plan and timeline.</li> </ul>

# Introduction to the review

The UK is internationally renowned for its longitudinal studies and the high quality data they generate to inform and assess policy. The Longitudinal Study of Young People in England (LSYPE) is a unique example of one of these studies. Launched in 2004, it has been a highly successful study generating large volumes of policy-relevant data for analysis. The study follows the transitions of a representative cohort of young people in England into adulthood in greater depth than any other existing data source. The Department for Education (DfE) anticipates the need to run another large scale longitudinal study of young people to help build on current knowledge of transitions from education in the teenage years into early adulthood.

This report is the outcome of a balanced and independent review of the first LSYPE cohort conducted by NatCen. The aim of this review is to suggest improvements to the design and make recommendations so that, should a valuable second cohort take place, it is designed to inform policy and meet the needs of all potential users. LSYPE is a unique study, in that it follows the transitions of a representative cohort of young people in England into adulthood. The development of a second cohort, based on the research and recommendations included in this review, would make the study even more valuable. Whether a second cohort takes place or not, this review also provides a reference for those setting up longitudinal studies with young people with regards to best practice.

While there are areas of overlap between LSYPE and other major cohort studies, LSYPE is the only major longitudinal study focusing on young people's experiences and entry into adulthood that covers the whole of England. LSYPE can therefore offer unique insights into the experiences of young people and the impact these experiences have on their later lives. As such, a second cohort would be an extremely valuable resource for future policy. In particular, if a second cohort was not commissioned, there would be considerable gaps in policy knowledge concerning young people's experiences of school, their relationships with their peers and families, their experiences of transitions into adulthood and work, their aspirations for the future, and, how these relate to their family backgrounds and socio-economic circumstances. As described in the following Chapters, other major cohort studies will not bridge these gaps. The report describes in detail why commissioning a single longitudinal study focusing solely on individual young people is likely to provide better value for money than commissioning a number of smaller cross-sectional studies.

This review combines a number of methodologies and has brought together a team of experts from NatCen including survey methodologists, senior researchers, statisticians

and analysts. Public Zone, specialists in marketing communication were also commissioned by NatCen to design a communication strategy that could support a second LSYPE cohort and help ensure its success. Research for the review also included gathering information with a broad range of LSYPE users, academics, Government Departments, fieldwork agencies and representatives from comparative longitudinal studies.

This review has been published following a recent change of Government. At the time of writing it is too early to clearly identify the priorities of the new Government and these policies are likely to impact on the content and analysis of a second LSYPE cohort, which should be borne in mind when considering the recommendations included here. However, some priorities of the coalition have been identified, including greater choice in service provision, support for disadvantaged groups and an interest in increasing social mobility. It will therefore be important for a second cohort to track and assess changes in young people's lives that may result from these policy developments. The ability of LSYPE to do this will be discussed in detail in this review.

## Objectives of the review

This report addresses a number of key objectives:

- Review how LSYPE has been used both within DfE and externally.
- Critically appraise the methodology for the first LSYPE cohort (including sample scheme, size and selection, and interview mode).
- Explore the potential for data linkages with other (administrative) datasets.
- Identify evidence gaps that a second cohort could uniquely fill.
- Make recommendations on DfE's analysis plan for a second cohort.
- Review how LSYPE has been funded across different Government Departments.
- Make recommendations for an effective communications strategy.

## Methodology

A variety of methods have been used to undertake this review, including desk research, quantitative analysis, depth interviews and focus groups. Discussions have been held with key people involved in LSYPE from DfE and the independent research consortium running the LSYPE, as well as with academics who have used LSYPE data. In order to inform the communications plan for a potential second cohort, a consultation among first LSYPE cohort members and other young people was also conducted. To protect the confidentiality of those interviewed, no names or organisations are referred to in this report, unless explicit permission has been given. The methods used in the review are described in more detail below.

## Literature reviews

Literature reviews were carried out on a number of topics, including for example:

- The use and effectiveness of different incentive strategies.
- The pros and cons of using different interview modes.
- Studies of young people that have included interviews with parents.
- Methods used on other longitudinal studies for increasing response rates, especially among 'hard to reach groups' (non-contacts and refusers).
- Methods for reviewing advances in technology and cultural changes to help maximise response.

## Analysis of first LSYPE cohort data

Analysis of the first LSYPE cohort data was carried out to inform the chapter on response rates. Study response rates were calculated for each Wave of LSYPE. This included a description of how response rates had changed over time and a comparison of respondents and non-respondents, using socio-demographic variables from the previous Wave.

## Telephone interviews

A small sample of the first LSYPE cohort members were interviewed by telephone. These interviews obtained feedback on the design of the first LSYPE cohort and the members' experiences of taking part in the survey. Interviews were carried out with both current respondents and those who had refused to take part at any stage (refusers). Refusers were defined as those young people who refused in person or by proxy (i.e. through parents or other household members) to the interviewer or the fieldwork headquarters, or those who made an appointment for an interview but didn't actually attend (broken appointment).

The telephone semi-structured interviews were audio recorded with respondents' consent. Interviews followed a pre-prepared topic guide and lasted around 20 minutes. Interviews were analysed using a content analysis approach based on Framework, an analytic tool developed by NatCen. A total of 12 young people aged 19/20 who had all taken part in LSYPE were interviewed: five current respondents; three short term refusers (i.e. refused for only one Wave of the study); and four longer term refusers (i.e. refused for more than one Wave). Table 1.1 outlines the characteristics of those interviewed.

**Table 1.1 Summary of characteristics of young people who were interviewed**

Sex	Respondent Type	Wave 5 mode	Ethnicity	Current activity
Female	Respondent	CAWI	White	Studying f/t
Female	Respondent	CATI	BME	Studying f/t, employed p/t
Female	Respondent	CAWI	White	Studying f/t
Male	Respondent	CAPI	White	Employed p/t, studying
Male	Respondent	CAWI <sup>†</sup>	White	Studying p/t
Female	Short Term Refuser	CATI	BME	Unemployed
Female	Short Term Refuser	CATI	White	Studying f/t
Female	Short Term Refuser	N/A	BME	Working f/t
Female	Longer Term Refuser	N/A	BME	Studying f/t
Male	Longer Term Refuser	N/A	BME	Unemployed
Male	Longer Term Refuser	N/A	White	Training scheme
Male	Longer Term Refuser	N/A	BME	Training scheme

\*CATI=Computer Assisted Telephone Interviewing; CAPI=Computer Assisted Personal Interviewing; CAWI=Computer Aided Web Interviewing

<sup>†</sup>Respondent thought that he had taken part by telephone.

Young people who had taken part in the survey were asked what motivated them to participate. Long term and short term refusers were asked why they dropped out of the survey and what would have encouraged them to continue to participate. All of the young people were asked their views on incentives and preferences for mode of interview. The timing and location of interviews and respondents' knowledge about confidentiality, consent and data linkage requests were also discussed.

## Workshops with young people

Public Zone held four one hour workshops with groups of 13-15 year olds in four schools of varying levels of achievement in the South East of England.<sup>1</sup>

- Group 1:** 10 boys (5x13yr olds and 5x15yr olds) in a single sex, average achieving comprehensive school.
- Group 2:** 12 girls aged 13-15 in a private, high achieving all girls' school.
- Group 3:** Mixed gender group of 13-15yr olds in an average achieving independent school.
- Group 4:** Mixed gender group of 13-14 year olds in a low achieving state school.

No workshop respondent had taken part in LSYPE. The aim of the workshops was to understand how best to develop a study which would appeal to this age group, and to create an effective communications strategy that could be adopted for a second LSYPE cohort. The following topics were discussed:

<sup>1</sup> This was an opportunistic sample rather than one that was designed to be representative of the population.

- What would make young people most likely to participate?
- The role that incentives may play in retention.
- Gauging reactions to communication approaches and materials.
- How important is the overall purpose of the study to young people?
- Parental involvement in the study, and how this might affect the young person's perception of the study and likelihood to participate.

## Depth interviews

12 semi-structured face to face and telephone interviews with key stakeholders who are, or have been, involved in the first LSYPE cohort were also undertaken. These interviews involved talking with current and former members of the research team at DfE responsible for LSYPE, senior staff within DfE, two of the three survey organisations who formed part of the original consortium undertaking the survey, LSYPE consortium members from other Government Departments, policy makers within DfE, and Government and non-Government data analysts.

The interviews covered a range of issues, with topic guides tailored to reflect different stakeholders<sup>2</sup>.

## User Group Seminar

A User Group Seminar was organised which provided the opportunity to consult users before making formal recommendations about sample size (as this depends on precision of the analysis required), discuss options for interview mode and other elements of survey design and to discuss evidence gaps that the second cohort could fill.

Stakeholders included representatives from DfE, GfK NOP Research, The Millennium Cohort Study (MCS), the Institute of Fiscal Studies (IFS), The Childhood Wellbeing Research Centre (Centre for Longitudinal Studies), Leeds University, TNS-BMRB, Bristol University, Institute of Education and the Research and Enterprise Service.

---

<sup>2</sup> Topic guides are available from [aleks.collingwood@natcen.ac.uk](mailto:aleks.collingwood@natcen.ac.uk)

The day was split into 3 sections<sup>3</sup>.

- The first section discussed population issues, data linkage, the funding of a second LSYPE cohort, whether to interview one or two parents in addition to the young person, and extra topics that a second cohort should include.
- The second section focussed on the sample design as well as the various options and recommendations for interview mode.
- The final section included the presentation of a recommended analysis plan followed by discussion. Public Zone then presented and discussed ideas around an effective communications strategy.

## Structure of the report

In the next Chapter, the LSYPE is introduced and its uniqueness (in terms of the data collected) is highlighted. The Chapter includes recommendations for a second cohort to collect additional information that other comparative studies do not include, followed by the types of analysis that could be carried out using a second cohort.

Chapters 3 to 5 look at the methodology of the current LSYPE cohort. The sample and fieldwork design are discussed in Chapter 3, followed by an analysis and discussion of response rates in Chapter 4, and then a discussion of mode of interview in Chapter 5. Each Chapter follows the same format. First the methodology of the first LSYPE cohort is described and evaluated. The pros and cons of the different approaches are then presented followed by recommendations for a second LSYPE cohort.

Chapter 6 discusses whether a second cohort should include an incentive for respondents and if so, what kind. This is in relation to the success of incentives used for the first LSYPE and the strategies of other, comparative studies. Chapter 7 covers who should be included in a second cohort, examining who was interviewed in the current LSYPE, and considers the pros and cons of interviewing these people using different modes of data collection.

Chapter 8 explores the advantages of data linkage when conducting large scale longitudinal surveys and identifies studies that would be useful to link to a second cohort. This is followed by the funding history of the first LSYPE cohort in Chapter 9, which identifies ways to make a second cohort more appealing to other Government departments, as well as discusses some of the pitfalls that need to be managed when pursuing cross-departmental funding.

Finally, before a summary of the recommendations in Chapter 11, Chapter 10 outlines a communications plan, with the aim of making a second cohort better resourced, more appealing to respondents, and making results of analysis more accessible to users and the general public.

---

<sup>3</sup> The User Group Seminar Agenda is available from [aleks.collingwood@natcen.ac.uk](mailto:aleks.collingwood@natcen.ac.uk)

## 2 Overview of LSYPE and rationale for a second cohort

This Chapter provides an outline of the first LSYPE cohort, including the uses to which the data have been put and the unique benefits of developing a second cohort. This sets the context for the ensuing Chapters which provide the basis for this review and the recommendations for a second cohort.

### 2.1 Overview of LSYPE

#### Objectives of LSYPE

Longitudinal research provides an understanding of social change, of the trajectories of individual life histories and of the dynamic processes that underlie social and economic life, which is not possible from research based on cross-sectional data. The development of longitudinal studies in the UK has also underpinned advances in social science methods and the understanding of major social changes and policy interventions. The UK has a wealth of longitudinal data from cohort studies such as the National Child Development Study (NCDS), the 1970 British Cohort Study (BCS70) and the Millennium Cohort Study (MCS).<sup>4</sup>

The main objectives of the first LSYPE cohort are to:

- Gather evidence about the transitions young people make from secondary education through to further and higher education or training into economic roles in early adulthood;
- Enhance the ability to monitor and evaluate the effects of existing policy and provide a strong information base for future policy development; and,
- Contextualise the implementation of new policies in terms of young people's lives.

#### Who runs LSYPE?

LSYPE is owned by the Department for Education (DfE) and co-funded by the Department for Business, Innovation and Skills (BIS) and the Department for Work and Pensions (DWP). It is carried out by TNS-BMRB Social Research, as the lead contractor, in consortium with GfK NOP and previously Ipsos MORI. A number of other organisations are subcontracted to assist on specific aspects of the project, for example the University of Southampton worked on the Missing Data Strategy, NatCen has been

---

<sup>4</sup> See <http://www.cls.ioe.ac.uk/> for further details

working on the data enhancement and GIDE/SDA have developed an interactive web tool iLSYPE, to assist with data availability and usability.

The first LSYPE cohort has also been subject to several Advisory Panels, the details of which can be found in Appendix A.

### **How long has the study been running and what stage is it at?**

LSYPE is one of the largest pieces of research ever undertaken by the DfE. The LSYPE started in 2004 (Wave 1) when more than 15,000 interviews were conducted with young people in Year 9 (aged 13-14) and their parents. Respondents have been re-contacted every year. Pupils from both the independent and the maintained school sectors are included. In the most recent wave (Wave 7, completed in summer 2010), approximately 9,500 interviews were held with the respondents then aged 19/20.

### **How has the study altered between Waves?**

Topics included in each Wave of LSYPE are tailored to collect information relevant to the young person's stage of life (e.g. whereas Wave 1 of LSYPE asked about reasons for Year 10 subject choices, the most recent Wave included questions on experiences of further education, Apprenticeships and employment).

In Waves 1-4, parents were interviewed as well as young people, but from Wave 5 onwards only young people were interviewed. There was been a minority ethnic boost in the original sample design, however a further boost was carried out at Wave 4 to improve on the original boost and counteract disproportionate attrition among these samples over time.

In Waves 1-4, the mode of data collection was face to face interviewing. In Waves 5-7, a sequential data collection design was implemented, with respondents first invited to complete a web questionnaire, moving to the telephone option and then to face to face if they refused, or were unable to complete via other methods.

### **What does the LSYPE questionnaire cover?**

The LSYPE questionnaires have a very broad coverage, looking at significant issues which affect the lives of young people and impact on their transitions and pathways into adulthood. See Appendix B for a detailed list of LSYPE topics.

Core household and demographic information is collected every year. All other topics are asked at different Waves or combinations of Waves depending on the age of the young person and policy interest, and include the following:

- The young person's attitude to school and involvement in education
- Extra-curricular classes
- Special educational needs and disability
- Parental expectations and aspirations
- Individual parent questions
- Family activities
- Parental relationship with young person and contact with services (self completion)
- Reasons young person does not live with natural parents
- Household responsibilities and resources (self completion)
- Risk factors (absences, truancy, police contact, bullying) (self completion)
- Ambitions for the future
- Friendships and socialising
- Family formation
- Higher education
- Employment

Questionnaires are available using the following web link

<https://ilsype.gide.net/workspaces/public/wiki/Questionnaires>. The highly routed and complex questionnaires allow researchers to capture the transitions young people are making and the wide ranging paths that they might follow.

## **2.2 How has LSYPE been used both within DfE and externally?**

The UK has an excellent international reputation for producing high quality longitudinal datasets which are used by analysts across the world, and LSYPE is a fine example of this. Data from Waves 1-6 are currently publicly available from the UK Data Archive<sup>5</sup> and, under some circumstances additional data is available from DfE directly. This availability has led to LSYPE being highly regarded by researchers, as evidenced by the extensive use of the data in the academic community as well as by a wider range of users. DfE invested in a data enhancement programme to ensure that the data are as useful and accessible as possible. This involves data cleaning, dealing with missing data and making the data more accessible to non-specialist users. In addition, in 2009 DfE launched 'iLSYPE'<sup>6</sup>, an online data exploration and management tool which aims to provide easier access to the data and documentation.

---

5 [www.data-archive.ac.uk/](http://www.data-archive.ac.uk/)

6 See: <https://ilsype.gide.net>

## How have the data been used?

The LSYPE data are regularly used within DfE to help inform policy development. For example, the data has been used to understand the characteristics and activity trajectories of young people who are Not in Education, Employment or Training (NEET); it has fed into the evidence base to raise the participation age; it has provided useful information on the characteristics of teenage parents, and, is currently being used to feed into the evidence base for the Special Educational Needs/Learners with Learning Difficulties Green Paper and the Schools White Paper. Data are also used by other Government Departments to inform policy development and by academics to investigate an array of associations between the measures available in the study.

In November 2008, DfE contracted its first Research and Analysis Framework Agreement to be used exclusively for LSYPE and Youth Cohort Study (YCS) data. By subcontracting analysis, the Framework allowed these datasets to be used much more extensively than DfE could manage in-house due to resource limitations. DfE has published three National Statistics Bulletins (in June 2008, June 2009 and July 2010) which illustrate how data from both the LSYPE and the YCS can provide detailed descriptions of the behaviours and experiences of young people and their families. The 2008 Bulletin focused on the activities of the cohort when aged 16, the 2009 Bulletin on respondents aged 17, and the 2010 Bulletin on respondents aged 18. All three Bulletins draw on LSYPE data to investigate trends over time.

The 2008 Bulletin can be accessed via:

<http://www.dcsf.gov.uk/rsgateway/DB/SBU/b000795/index.shtml>

The 2009 Bulletin can be accessed via:

<http://www.dcsf.gov.uk/rsgateway/DB/SBU/b000850/index.shtml>

The 2010 Bulletin can be accessed via:

<http://www.dcsf.gov.uk/rsgateway/DB/SBU/b000937/index.shtml>

## Published analyses

DfE has an ever-growing set of publications stemming from analysis of LSYPE data. Recent publications include topics such as bullying, school disengagement, risky behaviours and social activities, drivers and barriers to educational success, community cohesion, the use of information, advice and guidance, and alcohol consumption. LSYPE findings have also been published by other Government Departments, including the Cabinet Office, BIS, the Department for Environment, Food and Rural Affairs (DEFRA) and DWP. The literature review also found 19 publications that had been published by external organisations, on a wide range of topics. A full list of the publications can be found in Appendix C.

## External analysis being undertaken

A substantial amount of external research using LSYPE data has, and is, taking place. All data users who have downloaded data from the UK Data Archive were contacted directly as part of this review and asked to provide further details regarding the use to which the data is being put. Data users included University departments, professors, lecturers, postgraduate and undergraduate students, charities and independent research organisations. Much of this external data analysis is currently unpublished. A brief description of the content and analysis is provided in Appendix C.

## Other uses of LSYPE data

LSYPE data are used as study material within several university programmes, for example as part of the coursework for a Market Research Society Diploma. MRS is an awarding body for qualifications in market and social research. One of the units in their Level 7 Diploma qualification is 'Analysing and Interpreting Quantitative Market & Social Research Data'. For this unit, a dataset is selected from the UK Data Archive to form the basis of the candidate assessment. Examples of the questions from a current paper which has been set by the MRS are provided in Appendix C.

UK Data Archive users include students who have downloaded LSYPE datasets to use for assignments undertaken as part of their university course. For example, one student studying for an educational doctorate had downloaded the LSYPE data to look at the correlation between gender, age and the amount of homework pupils completed, as well as at the impact of modern technology (e.g. watching television, surfing the net, etc) on the amount of school work that young people complete at home.

Thus, the LSYPE has provided valuable data for a wide range of users, which is highly likely to continue should a second cohort occur. The data is widely used, which indicates a high demand and the creation of a second cohort will maintain and likely increase this demand. LSYPE is the only study following young people in such detail, particularly around their transitions to adulthood and the pathways through the teenage years. Exploiting this uniqueness ensures that the LSYPE provides clear value, as it can be put to myriad uses to inform policy and wider academic research agendas. Because of the LSYPE's wide ranging content it has a multitude of uses for Government policy and enhancing understanding of the role the Government play in the lives of young people.

Examples of policy areas where academic research using LSYPE has been carried out include differences in attainment between ethnic groups (British Educational Research Journal), young people's orientations and expectations for the future (Journal of Youth Studies), and the social structure of the 14-16 curriculum (International Studies in Sociology of Education). See Appendix C for further details. In addition, a lot of academic work using LSYPE data has been published in peer-reviewed journals,

thereby providing a further source of policy-relevant information without the need for specifically commissioned Government research.

### **2.3 The benefits of a second cohort and evidence gaps**

Despite understandable concerns regarding the expense and time required to undertake large longitudinal surveys, there is a strong argument that these studies make a unique, significant contribution to the national evidence base. Longitudinal information is crucial to understanding the impact of policy interventions on outcomes for young people, taking account of background information and prior experiences of the individuals concerned. Collecting this information in a large multi purpose survey is efficient, as it allows multiple policy interventions to be assessed. It allows a sufficient sample size to look at impact on important subgroups of the population, which could, for example, underpin Equality Reviews. It also allows analyses which may provide accounts of causality, by looking at the timing of interventions and changes of attitude and behaviour. A second cohort would also enable comparison with the first cohort allowing observation of changes between the two.

It is clearly important that a second LSYPE cohort is designed with consideration of the relationship it has to existing data sources. The particular strengths of the second LSYPE cohort (relative to the other cohort studies) must also be accurately and robustly described and opportunities for cross-study research be exploited. This section explores the content of other major longitudinal studies and assess the unique information that would be lost if a second LSYPE cohort was not commissioned.

#### **Content of the first LSYPE cohort and how this compares to other studies**

Unsurprisingly, past Waves of the first LSYPE cohort have focused on the educational experiences of young people, but other issues have also been covered including their views on local areas, community cohesion, participation in social activities, participation in risky behaviours, crime or anti-social behaviours, health, and their aspirations for the future. By including the parents or guardians in the survey the first LSYPE cohort has ensured there are data relating to the parents' involvement in the young person's education, as well as the socio-economic and demographic details of their household. Main topics for the parental survey include the young person's family background, parental socio-economic status, personal characteristics, attitudes, experiences and behaviours, attainment in education, parental employment, income and family environment as well as local deprivation, and the school(s) the young person attends/has attended. See Appendix B for a detailed list of questionnaire sections, topic lists for each Wave of the first LSYPE cohort.

As part of this review the main aims and topics covered by other key survey and administrative datasets, which cover the same age range as LSYPE, were examined.

The studies included were:

GUS	Growing Up in Scotland
FACS	The Families and Children Study
BCS70	British Cohort Study 1970
EPPSE	Effective Pre-School, Primary and Secondary Education (EPPSE 3-14)
ALSPAC	Avon Longitudinal Study of Parents and Children
US	Understanding Society
MCS	Millennium Cohort Study

In Appendix D the content of each study is presented highlighting the main topics covered, the distinct modules and any overlap with LSYPE. Four studies were investigated in more depth to compare areas of coverage with those being suggested for a second LSYPE cohort should one go ahead. These are discussed separately, below.

### Growing Up in Scotland (GUS)

GUS is a study that follows the lives of a national sample of Scotland's children from infancy through to their teens. This is one of the largest longitudinal studies ever done in Scotland and provides information that will help develop policies affecting children and their families in Scotland. The aim of GUS is to measure childcare, education, social work, health and social inclusion. GUS shares a number of similarities with LSYPE, it has a similar age group focus (although children are followed from an earlier age). However, in contrast to LSYPE, GUS has a wider research focus, investigating a number of areas in child development in greater depth than is intended for LSYPE. Many of these areas could be successfully incorporated into LSYPE through the integration of administrative datasets (particularly from health records). As a result, LSYPE would be able to place greater focus on areas related to pupils' educational achievement, and aspects of their home life and personal relationships. Furthermore, GUS is a Scottish study, and therefore not policy relevant to DfE, which has responsibility for young people in England.

### The Families and Children Study (FACS)

FACS is a panel study of approximately 7,000 families in Britain, investigating the circumstances of all families with dependent children. It provides nationally representative cross-sectional estimates for all households with dependent children as well as panel data for all six of the annual waves that have been completed thus far. FACS's major focus is on the economic circumstances of the family and on parents, rather than on young people *per se*. In contrast, LSYPE's major focus is on the transitions of young people and how different factors have promoted positive outcomes or led to situations where interventions may be necessary in order to improve outcomes. While FACS looks at current economic circumstances, LSYPE obtains a large amount of information related to how young people develop and how this leads to future economic

and social circumstances. In addition, FACS is now complete and will therefore not be a source of future data on young people and their transitions to adulthood.

### Understanding Society (US)

Understanding Society (US) is a large-scale study of the socio-economic circumstances and attitudes of up to 100,000 individuals in 40,000 UK households. The focus of US is on major societal trends and how they develop, as opposed to how young people per se develop into adulthood. US has a wide topic list and a particular focus on issues such as identity, beliefs and attitudes of individuals of all ages. DfE is interested in US, from the perspective of the issues affecting children and young people. However, this is not the main focus of the study and there is little depth in the sections of the study that focus on young people and little flexibility in the questions asked. In addition, US's age definition of 'youth' does not match that used by the DfE suggesting it may be of limited value to the Department.

### The Millennium Cohort Study (MCS)

MCS is a multi-disciplinary research project following the lives of around 19,000 children born in the UK in 2000/2001. It is the most recent of Britain's world-renowned national longitudinal birth cohort studies. MCS's field of enquiry covers such diverse topics as parenting; childcare; school choice; child behaviour and cognitive development; child and parental health; parents' employment and education; income and poverty; housing, neighbourhood and residential mobility; and social capital and ethnicity. As MCS is strong on socio-economic data, a second LSYPE focusing more on other areas such as attitudes and aspirations and life transitions would complement this study. As these two cohorts of young people would be almost the same age in 2012 (MCS respondents aged 12, LSYPE respondents aged 13/14), a second LSYPE cohort would need to be clearly differentiated from MCS. LSYPE has a strong focus on school factors, which could be expanded upon in a new cohort to further differentiate the study from other cohorts of a similar age. Another important advantage of a second LSYPE is that pupils would be clustered by school, enabling researchers to consider school effectiveness, which the MCS does not allow.

### Summary

In contrast to the first three examples, LSYPE only investigates young people, and focuses on major factors that are understood to have an important impact on their school experience and development into adulthood. Factors such as the impact of school and subject choice, relationship with parents and parents' attitudes towards their child's education, for example, are not investigated in Understanding Society but are crucial policy issues. LSYPE also uses a cluster sample of young people in schools, whereas Understanding Society and FACS are sampled at the household level. LSYPE is

therefore in a better position to examine the impact of different schools and types of schools.

This summary has shown that, while there are areas of overlap between LSYPE and other major cohort studies, LSYPE is the only major longitudinal study focusing on young people's experiences and entry into early adulthood, that covers the whole of England where it is also possible to examine young people's experiences at the level of the school. LSYPE can therefore offer unique insights into the experiences of young people and the impact these have on their later lives, and as such is an invaluable resource for policy.

If a second LSYPE cohort was not commissioned, there would be considerable gaps in policy knowledge concerning young people's experiences of school, their relationships with their peers and families, their experiences of transitions into adulthood and work, their aspirations for the future and how these relate to their family backgrounds and socio-economic circumstances. Other major cohort studies will not bridge these gaps. The section below describes in detail why commissioning a second cohort of a single longitudinal study focusing solely on individuals is likely to provide better value for money than commissioning several smaller cross-sectional studies.

## 2.4 Value for money

A second cohort is a long term and high cost commitment, involving complex design, fieldwork and data enhancement. It is therefore essential, particularly in the current financial climate and with increased pressure on budgets across Government, that such a study should be able to demonstrate value for money to potential funders. This section describes the main ways in which a second LSYPE cohort would provide good value for money and the potential pitfalls of not commissioning such a study.

### Maximising the value of the first LSYPE cohort

One of the key strengths a second cohort would provide is the opportunity for comparing the lives of young people in this cohort, with those in the first LSYPE cohort. This would provide an excellent evidence base of policy changes between the two cohorts and the impact they have had on the lives of young people. Moreover, it would make a study of this kind unique and valuable in providing insightful analysis as soon as the first Wave of data became available.

Such comparisons are likely to be particularly useful in light of the change of Government that has occurred in 2010. A second LSYPE cohort would remain in compulsory education or training longer than the first cohort, providing another basis for important comparisons between the two samples in line with plans to increase the participation age, which will increase the compulsory participation in education or training to age 18 by 2015. In addition, the first LSYPE cohort reached the end of compulsory

education in a time of relative economic prosperity, whereas a second cohort is likely to reach this milestone in a time of austerity. This difference between these two cohorts is likely to be large because of policy changes that will impact on them and the differences in the global fiscal climate. Running a second cohort would allow these differences to be studied in-depth, and to understand the impact of these changes.

A second cohort would add value to the first cohort and extend its useful lifetime in terms of analysis opportunities. Similarly, the first LSYPE cohort has the ability to add value to a second cohort by providing a wide base of data users who are already familiar with the structure and content of the LSYPE study. This user base can be easily translated to a second cohort, and should result in wide usage of LSYPE data from the outset. This pre-established network of data users would ensure that the data is analysed and used quickly (much faster than if a brand new longitudinal study was set up), making it much more valuable and timely.

### Meeting current Government priorities

The new coalition Government has outlined a number of key policies of relevance to children and young people, and a second LSYPE cohort would be instrumental in providing a sound evidence base for, and evaluating the success of, these policies. Such areas may include early intervention for children and young people at risk of multiple disadvantage, encouraging shared parenting, providing greater support to disabled children and those with special educational needs, protecting children in the event of family breakdown, introduction of free schools, the National Pupil Premium, increased powers being delivered to school teachers and a National Citizenship Service for 16 year olds<sup>7</sup>. A second cohort would have the ability to investigate the impact of these policy changes on young people as they move towards adulthood, as well as evaluating the effect of other policies not specifically targeted towards young people, but which may affect their parents or financial circumstances.

More generally, the emphasis placed on the Big Society – giving communities more powers and encouraging people to take an active role in their communities – and on improving social mobility, is likely to represent a change in the structure of young people's lives, whether at school, in work or at home. A second LSYPE cohort would therefore provide evidence to the DfE and other Government Departments on how these changes impact on young people. For example, a second LSYPE cohort would enable comparisons to be made in social mobility increasing with age. Existing data sources do not have the ability to make such comparisons and track young people over time as they grow older.

---

<sup>7</sup> [http://www.cabinetoffice.gov.uk/media/409088/pfg\\_coalition.pdf](http://www.cabinetoffice.gov.uk/media/409088/pfg_coalition.pdf)

## Providing economic value

A second LSYPE cohort would provide economic value by producing evidence to be used in the design of effective policies, including evaluations, cost-benefit analyses and value for money assessments. The survey is cost effective, because evidence to contribute to a large number of policy areas can be obtained from a single cohort and tracked over time, rather than requiring multiple smaller (but not necessarily cheaper), less wide-ranging studies to be commissioned.

The first LSYPE cohort has a broad range of measures available that capture the background and context of young people's lives. These measures enable researchers to go far beyond the kind of studies that are possible with administrative data alone (such as the National Pupil Database). For example, in an exploration of the unique contribution of individual characteristics to the differences seen in young people's attainment, and providing specific and valuable data not available via existing studies. Further, these measures enable researchers to explore the underlying processes that might explain how circumstances and behaviour impact on young people's attainment. For example, a number of recent studies using LSYPE data have explored behaviours such as bullying, alcohol consumption and other risky behaviours in the context of young people's family background, socio-economic circumstances, the areas they live in, relationships with their parents and attitudes and aspirations for the future. This type of contextual examination would not be possible using only administrative data.

As a longitudinal study following the same cohort of young people over time, a second cohort would make it possible to develop far more powerful analyses than could be achieved with a cross-sectional study. Analyses could be designed to explore individual change as well as the factors that contribute to that change, enabling answers to key questions, such as what leads young people who are similar in many ways to follow quite different pathways? Furthermore, a longitudinal design potentially enables the identification of causal relationships between events, because their chronological order is more accurately measured.

## Transparency

The new Government's focus on transparency can be served by LSYPE, as data from the first cohort are freely available (via the UK Data Archive<sup>8</sup>, iLSYPE<sup>9</sup> and direct from DfE) and concrete products of the financial investment can be seen. Data from a second cohort should also be available for users in the same way. There is strong evidence of demand for this kind of data, so it would be fully exploited once available.

---

<sup>8</sup> <http://www.data-archive.ac.uk/>

<sup>9</sup> <https://ilsype.gide.net/workspaces/public/wiki/Welcome>

## Additional topics that a second cohort could cover

Both the stakeholder interviews and the User Group Seminar agreed that the current content for LSYPE should remain and that there were a number of areas that a second cohort should cover and new topics that could be introduced, although it was acknowledged that practical constraints on questionnaire length affect how much can be covered. It was also generally agreed that if there are any gaps in LSYPE data then, if possible, it would be best to use administrative data to fill them to reduce study costs. It is inevitable, due to the limited capacity of an interview, limited numbers of administrative sources to link to, and the rapid policy changes which can occur, that there will be evidence gaps in a longitudinal study and this must be recognised and accepted. In such cases these gaps can be filled by cross-sectional data.

Suggested additional areas that could be included should a second cohort take place are summarised in Table 2.1 below. It must be noted that whilst these are all desirable areas, not all will be achievable given questionnaire length constraints and the balance of DfE's needs versus the requirements of academics/analysts. Some sections of the existing questionnaire would need to be dropped to make way for new topics and this would need to be balanced against comparability with the first LSYPE cohort.

**Table 2.1 Future topics for a second cohort as recommended by stakeholders and users**

Area	First LSYPE cohort coverage includes	Suggestions for a second cohort
Health	<ul style="list-style-type: none"> <li>• Disability/long term illness.</li> <li>• General health in last 12 months.</li> <li>• Birth weight.</li> <li>• Ability to perform everyday tasks and the effects on social life.</li> <li>• The extent to which a physical disability has inhibited daily activities.</li> </ul>	<ul style="list-style-type: none"> <li>• Height.</li> <li>• NHS numbers*.</li> <li>• Mortality data*.</li> <li>• Cancer registrations*.</li> <li>• Weight.</li> <li>• BMI.</li> <li>• Waist and hip measurements.</li> <li>• Physical activity and nutrition.</li> <li>• Blood pressure.</li> <li>• Self harming.</li> <li>• Eating disorders.</li> <li>• Relationships and sexuality.</li> <li>• Mental health.</li> <li>• General wellbeing.</li> </ul>
Psychological measures and personality measurement	None.	<ul style="list-style-type: none"> <li>• IQ or cognitive ability.</li> <li>• Personality traits, (e.g. Big Five personality traits).</li> <li>• Non-cognitive function scores.</li> <li>• Young people's thought processes about the future.</li> <li>• Perceptions of risk (behavioural economics).</li> <li>• Projected advantages of participating in positive activities.</li> <li>• Whether certain young people overestimate or underestimate risk.</li> <li>• Levels of confidence in making decisions (to look at behavioural biases).</li> <li>•</li> </ul>
Education	<ul style="list-style-type: none"> <li>• Attitudes to school.</li> <li>• Risk factors (including truancy, bullying).</li> <li>• Future plans and advice (including Connexions support and higher education aspirations).</li> <li>• Study support.</li> <li>• Rules and discipline at school.</li> <li>• Year 10 subject choices.</li> <li>• Reasons for Year 10 subject choices.</li> <li>• Subjects being studied (including favourite and least favourite).</li> <li>• Unfair treatment and discrimination.</li> <li>• Education Maintenance Allowance.</li> <li>• Absence from school (unauthorized).</li> <li>• School history.</li> <li>• Qualifications being studied for.</li> <li>• Higher education.</li> </ul>	<ul style="list-style-type: none"> <li>• Career aspirations.</li> <li>• Whether actively seeking career advice.</li> <li>• Parents' perception of and thoughts about the cost of education and risk of long term debt.</li> <li>• Information about the perpetrators of bullying.</li> <li>• Linked to nutrition, the possibility of including questions about school dinners and packed lunches to link this to free school meal information.</li> </ul>

Parenting	<ul style="list-style-type: none"> <li>• Parental/family engagement in education.</li> <li>• Whether parents talk about reports/GCSE choices etc with the young people.</li> <li>• Parental attendance at parent teacher evenings.</li> </ul>	<ul style="list-style-type: none"> <li>• Parenting style – how people parent teenagers.</li> <li>• Becoming a parent (i.e. looking at young people becoming parents themselves – an important transition to adulthood).</li> </ul>
Other areas	<ul style="list-style-type: none"> <li>• Influence of different organisations, role played by mentoring schemes and role of schools/parents/employers in motivating young people in education, work and community engagement.</li> <li>• More sibling information (birth order, age and sex of siblings, Sibling attainment and post-school destination).</li> <li>• More data on friendship groups (to build up a picture of the sort of young people the cohort member associates with).</li> <li>• Child maintenance (payments, receipt).*</li> <li>• Amount of contact with absent parents.</li> <li>• Transitions to work information related to attaining a consistent income of the young person.</li> </ul>	

\* Indicates data that could be obtained by linking to other data sources

Many Government policies are in place directly affecting young people in this age group and children approaching this age. It is important that a second cohort would collect data permitting an analysis of the impact of these policies. A number of these policies are summarised in Appendix E.

### Improvements on existing information collected

It is also vital to maintain a strong connection with the first LSYPE so as to not lose the benefit of comparison. However, the quality of some information currently collected in LSYPE should also be improved upon if a second cohort went ahead. For example, it would be essential to improve the quality of the information on family income that is currently collected, given the high volume of missing data and inconsistency of question wording across waves. There are also other examples where question wording changes between Waves, such as those relating to Information, Advice and Guidance, which makes it difficult to assess true change over time. The process of collecting this information needs to significantly improve. To help this happen, it is vital to try and maintain consistency between Waves in a second cohort more thoroughly than in the first. The next section outlines a range of research questions that a second LSYPE cohort could be used to answer, including an appropriate analysis strategy.

## 2.5 Analysis of a second cohort

The potential analysis plan set out here draws on the specific strengths of LSYPE and highlights examples of some of the insightful studies that could be achieved if a second cohort were to occur. The scope of an LSYPE is such that there are numerous research questions that could be addressed. To provide examples of this, suggestions have been made covering key thematic policy strands: young people's attainment; the impact of school choice; improving the achievement of disadvantaged young people; school curriculum; transitions into further and higher education, training or employment; vulnerable young people, young people's engagement in risky activities, and families, schools and communities. For each section a research question is posed followed by a brief description of the analysis. It is worth noting that this does not in any way represent an exhaustive list but for illustration only.

### Young people's attainment

Potential Research Question: *Which characteristics are important in predicting differences in attainment, and have these changed over time?*

Once the first Wave of a second cohort became available a relatively simple analysis could be carried out exploring differences in young people's attainment according to gender, ethnicity, religion, socio-economic background, parental education, housing tenure, level of income, family type and level of neighbourhood deprivation (controlling for all of these factors within the same model). There is already evidence highlighting the importance of individual characteristics and socio-demographic background on differences in young people's attainment. This research would add to the understanding of these relationships whilst taking into account the impact of policy changes and other social developments. The depth of information that LSYPE provides would also enable a very robust analysis relating to differences in young people's attainment beyond that achievable with administrative data. Furthermore, comparison with a similar analysis using the first LSYPE cohort would enable researchers to identify whether differences in attainment have changed over time, identifying, for example, which types of young people who were falling behind and in need of extra support.

The analysis could be extended to examine the influence of measures which reflect the context of young people's experiences within the family and school to see whether these contextual factors explain any of the differences found in young people's attainment. This could include the young person's own aspirations, parental aspirations and attitudes to education, parental discipline, parental supervision, parental closeness with the young person (including time spent together), whether or not they supervised their child's homework, levels of school discipline, quality of the young person's relationship with

teachers and extra curricular activities, to name just a few of the appropriate measures that would be available for this type of analysis.

## Decisions around school choice

Potential Research Question: *What are the factors that contribute to school choice and has this changed over time?*

A second cohort would provide opportunities for exploring change in the decision making process surrounding school choice which may result from current legislation to increase the options that are available to young people. Comparisons with the original LSYPE cohort would show whether the factors determining school choice (including the choice to opt for independent schools) have changed, and whether parents are following different avenues to obtain information on schools. Of interest would be whether patterns of behaviour are becoming more similar or different across certain groups of individuals, in particular, whether there is evidence of a growing equity in school choice among parents from different socio-economic backgrounds. In addition, it would be possible to explore whether parents behave differently depending on which part of England the young person lives and whether they live in an urban or rural area.

## Improving the achievement of disadvantaged young people

Potential Research Question: *What are the important factors contributing to increased achievement among children from disadvantaged backgrounds?*

Administrative data provides good information on differences in attainment and would be an obvious starting point for providing evidence on improvements in the achievement of disadvantaged young people. The added value of a second cohort would be in the provision of a more accurate representation of the relationship between disadvantage and attainment. Information on parental socio-economic position, family employment, family income, family type, parental education, tenure, and neighbourhood could be used to develop a far more complex and realistic account of young people's circumstances. In addition, this would enable researchers to identify which aspects of disadvantage (i.e. income, housing, or neighbourhood) are the most critical. The results of this analysis could then be compared with the findings from an equivalent study using the first LSYPE cohort data to see whether there had been any change in the achievement of disadvantaged young people relative to their peers.

Beyond comparative analysis, the study could be extended to explore underlying processes that explain how disadvantage leads to disparities in young people's attainment. This could be achieved by examining the link between disadvantage and experiences within the home (for example, parental aspirations and engagement in their child's education, parental attitudes to education, quality of the parent-child relationship,

parental discipline, supervision, time spent together, material wellbeing, family type, etc.) to see whether these factors could account for some or all of the relationship between disadvantage and attainment. Again, by comparing the results with a similar analysis on the first cohort it would be possible to see whether these processes have changed over time. It would be also possible to identify 'protective factors' that counteract the impact of disadvantage on achievement. Once further Waves of data become available, the study could be extended longitudinally to explore differences in the destinations of young people following post compulsory education and training.

## Vocational qualifications

Potential Research Question: *What is the impact of increasing vocational pathways for young people's engagement and attainment?*

In the advent of a second cohort a comparative study could be developed to examine change in the number of young people opting for vocational type pathways following the broadening of the school curriculum. Using a wide range of measures describing young people's individual characteristics and socio-demographic background the study could also identify the kinds of young people who follow vocational pathways, and assess whether this has changed over time. Analysis could explore the reasons that young people give for opting (or not) for vocational subjects and provide evidence on whether the provision of vocational type courses has an impact on the engagement of young people previously identified as disengaging, or at risk of disengaging, from education. This could be achieved by exploring changes in the attitudes, aspirations, behaviour and relative attainment of young people prior to and following the uptake of vocational type courses in Year 10 subject choices. As later Waves of data become available, which provide information on young people's experiences beyond compulsory schooling, the study could be extended to examine whether any positive impact identified is sustained through longer term engagement with education and training.

## Parental engagement

Potential Research Question: *What has been the impact of policies aimed at getting schools working with parents to encourage parental engagement?*

A key factor that has been related to young people's engagement and subsequent attainment is parental interest and engagement in their child's education. One of the strengths of LSYPE comes from the interviews with parents, which provide a broad range of useful data, which includes parental attitudes to education, aspirations for their child's education, how involved they feel in their child's education, as well as the quality of the relationship they have with the school (including the support they receive from the school).

A comparison of these measures across the first and second cohorts would provide insight into the success (or otherwise) of policy initiatives designed to develop positive relationships between parents and teachers, and to increase parental engagement and choice. The study could confirm the continued importance of parental aspirations and engagement by measuring their contribution to young people's attainment, aspirations and future destination, particularly among children from disadvantaged backgrounds. Furthermore, it is also possible to explore change in young people's aspirations and relate this to changes within the aspirations of their parents. That way it will be possible to get closer to identifying the causal direction of this relationship, i.e. do parent's changing aspirations impact on their child's aspirations, or are parents more likely to adapt their aspirations to their child's changing aspirations?

## Transitions into further education, training or employment

Potential Research Question: *How have young people's transitions into further education, training or employment changed over time?*

As a longitudinal study LSYPE could provide a very rich source of data for studying young people's transitions following post-compulsory education and training. A comparative study could be developed to explore the different pathways that young people follow in the context of new legislation to raise the school participation age to 18, providing strong evidence on the effect of the legislation. The analysis could examine differences in the stability of young people's educational, training or employment careers (i.e. whether they are less likely to 'drop out') to see whether these two additional years improve young people's life chances. It could also explore whether the effects are equally felt among girls and boys, different ethnic groups, young people from different socio-economic backgrounds and across different parts of the country. Preliminary analysis could be carried out even before the Waves of data documenting young people's transitions became available, exploring whether this has an impact young people's aspirations, attitudes to learning as well as their general future outlook, by comparing findings with the first LSYPE cohort.

## Vulnerable young people

Potential Research question: *How have the experiences of young people with special educational needs (SEN) improved over time?*

Following the release of the first Wave of data, analysis could be carried out exploring any changes in the experiences of young people with special educational needs, by comparing findings across both cohorts in terms of key outcomes, such as attainment, attitudes to school, aspirations, truancy, and engagement in both risky and positive activities. The strength of using LSYPE for this type of analysis is in the ability to control for a large number of potentially confounding factors that might otherwise account for

any differences found between young people with SEN, and other young people. This way there can be greater confidence that these differences (if they remain) can be attributed to having a special educational need, so that they can be addressed more appropriately in the future.

LSYPE also enables us to identify young people, whose family circumstances might make them vulnerable, perhaps because of poverty, living in a poor neighbourhood, living in care, living in an unstable household or having caring responsibilities. When multiple Waves of data become available, analyses could be extended to explore the differences in the transitions of these young people as they move out of secondary school and into further education, training, employment or NEET.

## Engagement in risky behaviours

Potential Research Question: *Are initiatives targeted at particular 'at risk' groups having the desired effect?*

Assuming similar measures of risky behaviours are recorded in a second cohort comparative analysis could be developed exploring changes in the prevalence of young people's engagement in truancy, smoking, drinking alcohol, cannabis use, graffitiing, vandalism, shop-lifting and fighting or causing a public disturbance. Using a broad range of demographic and individual characteristic measures it would also be possible to identify the characteristics of young people who are at greatest risk of engaging in these activities. A comparison of young people's experiences across the first LSYPE cohort and second cohort would then provide some insight into the success of initiatives designed to reduce or prevent risky behaviours among these groups.

Once further Waves of data became available, more complex and insightful analysis could be carried out to see whether there are particular events or 'triggers' that precede young people's engagement in risky behaviour. In addition it would also be possible to identify 'protective factors', termed so because they reduce the risk of engaging in risky behaviours among those who are (because of their circumstances) more likely to engage in these activities. Again, comparative research with the first cohort could identify whether the underlying processes which lead to engagement in risky behaviours have remained stable over time. Further still, because LSYPE is a longitudinal study it would be possible to employ very advanced analysis techniques enabling us to distinguish between young people whose delinquent behaviour is likely to be limited to their teenage years, from those likely to become 'life course persistent' offenders.

## Families, schools and communities

Potential Research Question: *What is the role of the family/school/area in helping young people develop and achieve?*

LSYPE is a multi-level or cluster sample, in which young people are sampled within schools, and these schools are themselves sampled within local authorities. It is also recommended that a second cohort should follow this same strategy. The benefit of this sampling approach is that (as well as contributing to lower survey costs) it enables the application of an analytical technique that helps identify the relative importance of different spheres of the young person's life to their development. This technique, called multi-level analysis, allows researchers to calculate how much of the variation in young peoples experiences are to do with their own and their families' characteristics, the schools which they attend, or the area in which they live.

All of the analysis examples outlined above could be carried out within this multi-level framework to enable a much richer understanding of how individual characteristics, families, schools and the local area contribute to young people's experiences, including their attainment and aspirations, the choices they make, their transitions into further education, training or employment, and their engagement in risky behaviours. Furthermore, by drawing comparisons between the first LSYPE cohort and a second cohort, it would be possible to map any change in their relative influence over time.

### 2.6 Advisory Panels

As part of this review the role that an expert Advisory Panel has played in the first LSYPE cohort and could likewise play in a second cohort has been explored. Findings from this are summarised here. This section also looks at the idea of setting up a separate Panel of young people to work in an advisory capacity.

#### The Previous LSYPE Advisory Panels

Seven Advisory Panels were established when the LSYPE project first began. Each Panel had a convener, and members comprised stakeholders from British and European universities and organisations as well as from those within the USA. The Panels made recommendations identifying variables that LSYPE needed to measure. Methods of measuring these were then presented at a consultative conference, and this formed the base of the LSYPE content. These Panels were only active at the initial set up of the LSYPE.

The seven topics were:

- School policy and curriculum
- Ethnicity
- Adolescent development
- Family and health
- Labour market
- Youth policy and practice
- Methodology.

## Recruiting stakeholders into an Expert Advisory Panel

Interviews with stakeholders showed unanimous agreement that that an Expert Advisory Panel would be useful for a second cohort. This would provide guidance on the issues described above and represent the interests of a range of stakeholders in the study. It is therefore advised that such an Expert Advisory Panel should be set up and consulted at the beginning of the study as was the case with first cohort of LSYPE. After the initial setup and consultation it would be useful to consult with the Expert Advisory Panel at timely intervals (at least annually).

A wider discussion needs to take place about who would be best to act in an advisory capacity for a second cohort. However, if it is possible to run multiple Advisory Panels with different types of interested parties alongside one another then this would be preferable. See Appendix F for a list of other cohort studies that have had Advisory Panels consisting of academics and/or policy experts.

## Recruiting LSYPE cohort members into a Young Person Advisory Panel

In addition to an Expert Advisory Panel, it is recommended that a second panel consisting of young people is also used. Such panels have been used by other studies to help guide the research and help the target population become more involved in the study rather than simply as research subjects. However, including such a 'Young Person Advisory Panel' would have to be carefully thought out and would be expensive to adopt. Young panellists would require extensive support and training to ensure that they have the skills and confidence necessary to meaningfully engage in the panel. However, if this need for support is adequately addressed, a panel of this sort could provide many benefits to the study. These include:

- Ensuring that the design, delivery and dissemination of the research is as relevant and accessible to the age group of the cohort as much as possible;
- Adding transparency to the decision making process within the study - young people can help guide the project and by doing so can learn about the study, understand why Government conducts such research and be involved in the decision making; and

- Increasing the credibility of the research among other young people and wider stakeholders - panel members can act as the voice for young people in the study, usefully promoting the study and encouraging involvement among their peers thereby maintaining response rates.

Whereas the Expert Advisory Panel would feed into the study using their expert knowledge and research experience, the Young Person Advisory Panel would feed in their expertise of being a young person. They can add a valuable young person perspective and can be used for quality assuring all materials produced for the study including the content of the questionnaire and incentives, to ensure that they are suitable for the intended audience. This could include the pre-test of questionnaires prior to wider piloting. The panel could also assist with the communication plan by designing or commenting on how messages from the study are disseminated to policy makers, young people and wider stakeholders.

Ideally, the Young Person Advisory Panel should consist of panel members who are approximately the same age as the study respondents they would be representing. They could be drawn from the same sample frame as that used for the study respondents. However, because of issues of respondent confidentiality, care should be taken not to select pupils from the same school. Parental consent would be needed for young people under the age of 16 to be involved in the Young Person Advisory Panel and appropriate methods of recruiting and maintaining the Panel would need to be discussed and agreed with DfE.

Engaging young people in this sort of activity can be challenging so consideration would need to be given to the mode by which they are engaged. At the point that the panel is set up, a series of regional face-to-face briefings could be arranged for panellists. The purpose of these initial meetings would be to define the Terms of Reference for the panel and provide training that empowers the young panellists to participate fully. Given that the panel would ideally be recruited of young people from a range of backgrounds, this training is likely to need to be quite bespoke and, in some cases quite intensive. These initial briefing sessions would need careful thought and should be run by experienced trainers and facilitators. Consideration should also be given to appointing a team of mentors whose job it would be to maintain contact with the panellists through a variety of mediums (including telephone, text and online) and ensure that the panellists both understand what is expected of them and have the skills and confidence to participate in the panel in the way they want.

Panel members should be incentivised to participate in the panels and, if travel was required, expenses for this should be met by the study's budget. Following the initial briefing sessions, the main mode of communication for the panel could be online. This could be done through a dedicated website or forum. While this mode of communication

is likely to encourage more engagement from the panel than face-to-face meetings alone, given the age of the panel members attrition is still highly likely. Steps should be taken to ensure the frequent replenishment of panel members, perhaps with new members recruited annually. Each new intake of panellists would need to be accompanied by a face-to-face briefing session for the new recruits.

If budgets for the study do not allow for a fully dedicated Young Person Advisory Panel to be convened, it may be possible to draw on existing DfE panels for this purpose, and this option should be explored at set up of the study.

## Areas for Consultation

It is important to gain information from both Advisory Panels about a variety of issues such as questionnaire content, the use of incentives and methods of keeping in touch with respondents in order to minimise attrition and to make sure that a second cohort serves the interests of the young people it represents. It would also be important to discuss the mixed mode design and how this will work in practice. The Advisory Panels could also be usefully consulted on the best ways to present literature on the LSYPE, in order to have the best possible impact on the young people and their parents in terms of their willingness to take part and understanding of the study. These views would enhance a second cohort throughout its process and could add to increased response rates.

## 2.7 Recommendations

In the previous sections ideas for the content of a second cohort and some examples of potential analysis of these data have been outlined. The recommendations from these sections along with those for advisory panels are summarised below:

### Content

- A key benefit of a second LSYPE cohort is the possibilities it would offer for comparison with the first LSYPE. It is therefore recommended that the majority of the questions from the first cohort remain in place for a second cohort.
- It is recommended that the following content be added to the LSYPE: further measures on health; psychological and personality measurements; further questions on school choice; career aspiration; parenting; social network information (friends, siblings, contact with absent parents, etc.); and sexual orientation. NatCen recommends that these areas should be incorporated into a second cohort if resources allow, and that more should be made of the opportunity to link to other datasets, if this represents a more appropriate way of collecting this data (see also Chapter Eight on Data Linkage).

## Analysis plan

- Analysis of a second cohort should focus on its particular strengths, drawing out what makes this study unique from other data sources. One of the most obvious strengths would be the possibilities it offers for comparative studies with the first LSYPE cohort, and NatCen therefore recommends that this should be a major focus of analysis work.
- Using the wealth of data collected via two LSYPE cohorts, studies could be developed to explore change in the experience and context of young people's lives and enable (at least to some extent) an evaluation of Government policies during the intervening period.
- The broad range of contextual information that is available on young people's lives could be used to develop a much better understanding of the differences in young people's experiences as well as the factors that contribute to this difference.
- Once later Waves of data become available, longitudinal analysis should be used to explore young people's development over time to gain a better understanding of the different pathways they follow.
- The exact content of the thematic and policy relevant questions that could be answered using a second cohort should be agreed in discussion with the DfE but is likely to cover thematic policy strands such as: young people's attainment; school choice; school curriculum; improving the achievement of disadvantaged young people; parental engagement; transitions into further education, training or employment; vulnerable young people, young people's engagement in risky activities, and the role of the Big Society (teachers, parents, communities) in helping young people develop. Of course, these themes could be adapted to take account of change in policy priorities.

## Advisory Panel

- NatCen recommends that two Advisory Panels are set up for a second LSYPE cohort. An Expert Advisory Panel consisting of stakeholders including academics and policy makers, and an Advisory Panel consisting of young people. Whereas the Expert Advisory Panel would feed into the study using their expert knowledge and research experience, the Young Person Advisory Panel would feed in their expertise of being a young person.

- Advisory Panels should be consulted separately, and at the setting up stage of a second cohort. Both Advisory Panels should then meet at timely intervals (at least annually) while the cohort becomes embedded. As attrition is highly likely in regards to a Young Person Advisory Panel steps should be taken to recruit new members to this group annually. It is suggested that the Advisory Panels should meet prior to the questionnaires being finalised.
- Advisory Panel members should be consulted on issues such as questionnaire content, design of questions, question placement and frequency, use of incentives, methods of keeping in touch with respondents, the design of the study and accessible ways of presenting study literature.

## 3 Sample and fieldwork design

This Chapter describes the sampling strategy used in the LSYPE, and describes possible improvements that could be considered for a second cohort. The first section outlines the sample of the first LSYPE cohort and summarises what worked well. Section 3.2 considers issues of defining the target and study populations, and 3.3 discusses choosing a suitable sampling frame. This is followed with a discussion of choice of sample size, the desirability of a boost of important subgroups and the issue of stratification. The final section focuses on the fieldwork design, including the ideal age of a new cohort, the year that fieldwork should begin, before concluding with recommendations that could be incorporated into the planning of a second cohort.

### 3.1 Strengths and weaknesses of LSYPE sample

As outlined in Chapter Two, the existing LSYPE is a large scale longitudinal survey of young people who started to be interviewed when in Year 9 or equivalent. Pupils were selected from those attending maintained schools, independent schools and pupil referral units.

Pupils were sampled using a two stage design. Schools were sampled at the first stage then pupils were sampled from each of the selected schools at the second. In maintained schools, selection probabilities varied so that schools were over-sampled on the basis of their deprivation status (as measured by the proportion of pupils receiving Free School Meals (FSM)); and pupils from some minority ethnic groups were over-sampled to ensure sample sizes were large enough to enable subgroup analyses.

Excluded from the survey were young people not on the school roll, boarders, children residing in the UK solely for educational purposes, and pupils in very small schools (those with fewer than 10 (maintained sector) or 6 (independent sector) pupils).

The LSYPE had several features that users suggested they would like to have retained should a second cohort be conducted. The sample size was regarded as adequate (though some users would have liked a larger study as subgroup analysis became difficult in later Waves, particularly in terms of some minority ethnic groups and disadvantaged groups, for example those with disabilities). The over-sampling of minority ethnic groups was regarded as a particularly useful feature of the survey, as was the over-representation of pupils eligible for FSM. Users also appreciated the two stage nature of the design, as this meant multi-level modelling was possible allowing analysts to estimate school effects on child outcomes. This would not be possible without the two stage approach.

Very few users commented on weaknesses in the sample design for the first LSYPE cohort. One criticism of the first cohort was that rural schools were not over-sampled. Some users expressed an interest in comparing rural and urban schools and queried whether a rural boost would have made this comparison more powerful. Another criticism was the method used to over-sample deprived students. Pupils eligible for FSM were over-represented in the sample because deprived schools had a higher selection probability. It was pointed out that it would have been preferable to over-sample pupils eligible for FSM directly rather than over-sampling their schools. However, these criticisms were considered to be minor and the sample design of the LSYPE was considered to be one of its strengths.

### 3.2 The target and study populations and exclusions

Should a second cohort take place it is likely that the young people of interest would be similar to those from the first LSYPE cohort, and that most Year 9 or equivalent pupils living in England should be part of the target population. Some Year 9 pupils would *not* be included in the target population, for example, those residing in the UK solely for educational purposes. The eligibility of these pupils would typically not be known until after they were sampled. They would therefore remain on the sampling frame with a positive probability of selection, but any such pupils selected would be screened out and recorded as ineligible.

It is likely that the study population – i.e. those pupils from whom selection is made – would differ slightly from the target population. This would arise in circumstances where schools or pupils are excluded from the sample because it is either too expensive or too difficult to contact them. The main candidates for exclusion would be:

- Pupils in small schools: The first cohort did not sample pupils in small schools, and stakeholders pointed out that this contributed to rural schools being under-represented. However, practical considerations made it uneconomical to include them. Excluding schools with 10 or fewer Year 9 pupils on the Schools Census excludes only 0.5% of all pupils in England, so it would not introduce any large biases into pupil-level analysis. However, it does exclude more than 10% of schools, and this might be unsatisfactory for performing school-level analysis. There is a demand for school-level analysis. It would be possible to include such schools in a second cohort (by under-sampling them), though there would be cost implications to this. This is explored in more detail in the clustering section below.
- Children not on the school roll: should a second cohort take place it is recommended that it follow the first LSYPE cohort, and exclude children not on the school roll. This is because there is no obvious cost effective way to sample them and they constitute only a small minority of the target population.

- Pupils whose addresses are missing on the Schools Census: The Youth Cohort Study (which also sampled from the Annual Schools Census as is suggested for the second LSYPE cohort – see below) excluded some pupils whose addresses on the Census were incorrect or incomplete, and it is likely that this would also be necessary, should a second cohort take place.

### 3.3 Sampling frames and clustering

For a second cohort, it would be advantageous to use separate sampling frames for maintained and independent schools due to the different sources of information held about these schools. These are discussed separately below.

#### Maintained Schools

Pupils in maintained schools could be sampled using the Annual Schools Census. This is a natural sampling frame for maintained schools, as it contains a rich amount of data to aid sampling and it is also possible to merge Key Stage 3 attainment data onto it.

The first LSYPE cohort was not able to sample directly from this database. Instead, a sample of schools was chosen and each selected school was contacted and asked to take part in the survey. If the school was willing to cooperate, a sample of pupils was chosen from within each of the cooperating schools. School cooperation was essential to gain pupil addresses. The Annual Schools Census now contains address information, so it is no longer essential to sample pupils through schools. It is recommended that should another LSYPE cohort occur pupils are selected directly from the Annual Schools Census. This would not only be more efficient, but would eliminate non-response at the school-level. However, it is regarded as good practice to inform schools of the study, should potential respondents have any queries.

Should a second cohort take place it is recommended that the first Wave of data collection would be face to face (see Chapter Five on mixed modes, for further details). To do this cost effectively a clustered sample is recommended (to reduce travel costs). Two obvious clustering methods would be to:

1. use schools as clusters; or,
2. use a geographical variable such as postcode sector or super-output area.

Making a decision on which is the more appropriate clustering variable - schools or a geographical unit is not clear cut. The main arguments of each are summarised in Table 3.1 below. On balance, the arguments for choosing schools as the Primary Sampling Unit (PSU) seem stronger. The main reason is that users are interested in multi-level modelling, which requires the school to be the PSU. Multi-level modelling is increasingly a leading method of choice for many researchers, and is likely to be increasingly so in the future. It is particularly useful for analysing educational data, as it allows for

characteristics of both individual pupils and the schools they attend to be analysed at different levels to determine the relative importance of school-level and individual-level effects on outcomes. It will therefore be useful for a second LSYPE cohort to facilitate this type of analysis in its design.

**Table 3.1 Advantages of using schools or geographical areas as the clustering variable**

Advantages of using schools	Advantages of using geographical areas
<ul style="list-style-type: none"> <li>• Ensures a minimum sample size will be obtained in each of the selected schools. This will allow for multi-level modelling.</li> </ul>	<ul style="list-style-type: none"> <li>• No need to exclude or under-represent small schools – all pupils can be included.</li> </ul>
<ul style="list-style-type: none"> <li>• Design would be consistent with first LSYPE cohort. This was mentioned as an advantage by some users, but is not regarded as a high priority.</li> </ul>	<ul style="list-style-type: none"> <li>• Smaller design effects and a more efficient design.</li> </ul>
<ul style="list-style-type: none"> <li>• School support could help with sample retention.</li> </ul>	
<ul style="list-style-type: none"> <li>• School support could help where address information is missing from the census.</li> </ul>	
<ul style="list-style-type: none"> <li>• This would mean using the same approach for maintained and independent schools. This is also not regarded as a high priority, but some users have mentioned it as being desirable.</li> </ul>	

A major disadvantage of using schools rather than geographical areas as the clustering variable is the cost of sampling pupils from small schools. In the first LSYPE cohort this was dealt with by excluding the smallest schools. This would also be a solution should a second cohort take place, but there are alternatives which could be considered. For example, it would be possible to take a sample of small schools (e.g. those with 10 or fewer Year 9 pupils), but set the selection probabilities to be relatively low. This would mean that some small schools would be selected, but fewer than would have been selected if an equal probability sample were chosen. The low selection probabilities would mean that these schools were under-represented, but the data could then be weighted to accurately represent the population. The achieved sample of pupils from these small schools would *not* be large enough to analyse in its own right; these pupils would be included simply to improve the analysis of the national sample (or samples of other subgroups). Including these schools would, however, introduce an additional cost to a second cohort due to the small cluster size. Despite this, NatCen would recommend adding a sample of small schools with a low selection probability in order to make sure

this population of pupils (who may have different characteristics from those who attend larger schools) is represented.

A second disadvantage of using schools as the clustering variable is that it is more expensive than using geographical areas because fieldwork assignments cannot be efficiently clustered – they tend to be spread over a wide area. Despite these two disadvantages, for the reasons given above, NatCen believes there are stronger arguments for choosing schools, rather than a geographical unit, as the PSU.

Once clusters are selected, a random sample of pupils from each cluster could be chosen. In the first LSYPE cohort, approximately 30 pupils per school were chosen and stakeholders felt that this was an appropriate size. NatCen therefore recommends that should a second cohort take place – and should schools be chosen as the clustering unit - this is the approach taken.

## Independent schools

The DfE does not collect pupil-level data from Registered Independent schools. However, independent schools complete a school-level Annual Schools Census for Registered Independent schools. This is a paper form collection held once a year in January and this can be used to sample independent schools. A second cohort could follow the first LSYPE cohort by first sampling independent schools, then taking a sample of pupils from within each of the selected schools. Selected schools would need to be contacted and asked to take part in the survey. If they were willing to cooperate, a sample of pupils could be chosen from within each of the cooperating schools. If the school did not cooperate then it would not be possible to sample these pupils, so school non-response would affect the sample. This could be dealt with when weighting the data.

### 3.4 Boosts and sample sizes

There was broad agreement in the stakeholder interviews that to maximise the value of LSYPE, a second cohort should provide sufficient numbers in key subgroups of interest to allow meaningful analysis, though which subgroups should be boosted was a topic of some debate. Boosting any group will improve the analysis of that subgroup, but leads to inefficiencies when performing an *overall* analysis. Because of this it is important to restrict the number of boosted groups in studies such as this. The groups to boost should be chosen after considering the type of analysis planned and the likely achieved sample sizes for the groups of interest, although such detailed plans are often difficult to achieve in longitudinal surveys as the possibilities for analysis are so broad. Although some stakeholders suggested boosts for small groups, such as teenage mothers, there would be little point in boosting such small groups as the numbers would still be unlikely to yield a sufficient sample size for robust analysis.

The final decision on which groups to boost would be made by the DfE. However, promising suggestions include:

- Minority ethnic groups, as in the first LSYPE cohort.
- Disadvantaged groups, as identified by an indicator such as free school meals (FSM) eligibility, the Index of Multiple Deprivation (IMD) or the Income Deprivation Affecting Children Index (IDACI).
- Lower-achievers as measured by KS3 results.
- Pupils in academies (or other types of schools).
- Schools in rural areas.

NatCen would recommend boosting minority ethnic groups, some disadvantaged groups (probably using the IDACI measure since FSM eligibility is likely to change with policy changes) and schools in rural areas. However, the number and types of groups to be boosted will also depend on levels of funding.

If pupils in academies (or any other type of school) were boosted, the boost should be at school-level. For example, academies could be chosen with a higher probability than other schools. Once a school was chosen, an additional boost could be selected on the basis of the pupil's characteristics (ethnicity, FSM status, KS3 results etc). This differs from the first LSYPE cohort, where *schools* with a high proportion of pupils eligible for FSM were boosted; for a second cohort directly boosting *pupils* eligible for FSM may be more robust.

The overwhelming view from users of the first LSYPE cohort was that sample sizes on a future LSYPE should ensure robust analyses for examining important subgroups. This was regarded by some stakeholders to be more important than the overall analysis. The first LSYPE cohort was designed to achieve a minimum Wave 1 sample size for each of the main minority ethnic groups. Rather than specifying a required sample size per subgroup, a second cohort could be designed to ensure a minimum *effective* sample size<sup>10</sup> for each important subgroup. Moreover, this should not be specified only for Wave 1, but for subsequent Waves as well. For example, to achieve an effective sample size of 600 Black African pupils in Wave 4 about 1,397 Wave 1 interviews with this group would need to be achieved. This would involve an issued Wave 1 sample of almost 1,900 Black African pupils. The calculations, and the response rate assumptions underlying them, are based on data from the first LSYPE cohort and the exact numbers would depend on the level of precision required by the DfE. Nevertheless, it is anticipated that an achieved sample size of at least 1,000 pupils in each of the main minority ethnic groups in Wave 1 would be required. This suggests that the overall Wave 1 sample size for a second cohort should be as large as that obtained in the first LSYPE

---

<sup>10</sup> The effective sample size takes into account the complex sample design, including the weighting, stratification and clustering. It measures the size of an (unweighted) simple random sample that would have provided the same precision (standard error) as the design being implemented.

cohort. Because minority ethnic groups are likely to be the smallest subgroups of interest (with only very small percentages of young people in many of the groups), it is unlikely that a decision to boost other subgroups (for example low achievers) would substantially increase this required sample size.

### 3.5 Stratification

If selecting maintained schools for a second cohort, stratification would be possible both at school-level and at pupil-level. Final decisions on which stratification variables to use would have to be made following discussions of which subgroups are to be boosted, as these decisions are interdependent. However, it is recommended that for maintained schools the following school-level stratifiers be considered:

- Type of school
- Regions (Government Office Regions)
- An urban/rural indicator
- The academic performance of the school (such as the proportion of pupils obtaining five or more GCSEs at grade C or above)

At pupil-level the following could be useful stratifiers:

- Ethnic group
- Gender
- KS3 performance or Deprivation indicator

Independent schools could be stratified on the basis of variables such as:

- Whether it is a single or mixed sex school
- Government Office Region (GOR)
- The academic performance of the school (such as the proportion of pupils obtaining five or more GCSEs at grade C or above)

Having considered the sampling strategy for a second cohort, the next section turns to the fieldwork strategy that may be adopted.

### 3.6 Age of cohort and fieldwork strategy

A considerable amount of change could be anticipated for pupils aged 14-19 between 2010 and 2015. For example, it is anticipated that by 2013, young people from the age of 14 onwards would be able to choose between pursuing general qualifications, and new, employer-designed 'specialised diplomas'. In line with this reform, in 2013 all young people in England are to be required to continue in education or training up to the age of 17 and by 2015 to continue in education or training up to the age of 18<sup>11</sup>. This change means that those young people who started Year 9 in September 2010 will

---

<sup>11</sup> The legislation for increasing the participation age was contained in the Education and Skills Act 2008.

continue in education or training to 17, and those in Year 8 (who started secondary school in September 2009) will continue until they are 18.

Two options which could provide evidence on the implementation of this change include surveying the first cohort of young people affected (those currently in Year 9 who will continue in education and training to age 17) or the second cohort affected (those currently in Year 8, who will continue in education and training to age 18). Only the second cohort would experience the full measure of remaining in education or training until the age of 18 however, and findings from this cohort are likely to be more robust for the formulation of future policy.

When this review was commissioned there was a very strong steer that the new cohort should start in 2012 to coincide with the second year of pupils being obliged to remain in education and training until at least 18<sup>12</sup>. The advantage of this is that it will accurately measure the impact these changes have on educational progress and transitions by capturing a full cohort of young people experiencing remaining in education until 18. However, given changes in Government since the study was commissioned, there might be opportunity to review this decision and consider whether or not a second cohort could start earlier. As young people are already forming ambitions and attitudes much earlier there is an argument to start as early as possible.

A major design consideration for a second cohort would be comparability with the first LSYPE cohort. By starting with young people of the same year group, direct comparisons could be made between the two cohorts in a very short space of time. As soon as the first annual collection of data had been attained and was ready to analyse any apparent differences in educational progress (of 13/14 year olds) could then be directly analysed in the context of Government policy changes that have taken place over that time period, for example.

In planning a second cohort it would also be important to note when other studies' fieldwork is taking place to make sure there are no major clashes which may impact on the study's success. If other large scale studies are in the field at the same time as a possible second cohort this might restrict the pool of fieldworkers available and therefore restrict options for who (which research organisations) can undertake a second cohort on behalf of DfE. There is also the issue of respondent burden as households may be participating in more than one study. Using 2012 as an example as where these overlaps may occur, it is predicted that the 2012 Birth Cohort Study may start then if this study still goes ahead, and the established Birth Cohort Studies (BCS70 and MCS) will

---

<sup>12</sup> It must be noted that the primary focus of a second LSYPE cohort would not be to measure the impact of this Act. However, it will indeed be a unique and invaluable way to do this.

be in field this same year. Details of the fieldwork timings of other relevant studies are in Appendix H.

NatCen therefore recommends that if a new cohort was to start in 2012 then it should begin when the young people are in Year 9 – to promote comparability with the first cohort – and when fieldwork commences requires careful consideration to ensure that the study adequately captures future policy priorities. If the fieldwork year changes, to 2013 for example, then the cohort should be in Year 9 in 2013.

## Fieldwork timing

The first LSYPE cohort fieldwork currently takes place in late spring and through summer on an annual basis. This timing was originally due to the expediency of being able to interview cohort members and their parents at home and without clashing with exam periods. Fieldwork timings have since remained the same so that cohort members know at what time of year they will be asked to take part. Discussions with the current fieldwork agency revealed particularly good responses when the fieldwork period runs over the holiday periods.

LSYPE respondents that were interviewed as part of this review were asked about the ideal timing and location of interviews. The current timing of fieldwork was felt to be satisfactory, since it did not clash with exams. Respondents felt that not coinciding with exams was important for the survey, and NatCen would agree, that such a clash is to be avoided if possible. NatCen would therefore recommend that the current fieldwork timings of spring and summer should be retained, but that these may need to be modified in exam years if they are likely to coincide with exam periods.

Views on where the interviews should take place (at home or at school), specifically in the earlier Waves of the survey, were mixed. Home was preferred when respondents felt comfortable and 'safe' in this environment and were happy to answer questions on more sensitive topics either within earshot of their family or in a space they felt sure was private. School was preferred when respondents had felt that it would provide a more 'private' setting than home. This reflected respondents' unease about family members being able to overhear answers to questions on sensitive topics such as alcohol consumption. However, since the most sensitive questions are asked in the self completion section of the questionnaire this is unlikely to be a substantive problem. NatCen therefore recommends that if interviewing face to face for the first three Waves, the ideal location for interviews is at home. Home is also the best option for interviewing parents. If mixed mode interviewing is introduced at later Waves by web/telephone then flexibility of location could be an option for those still having face to face interviews however NatCen would still recommend home as the ideal interview location for those respondents who were continuing to be interviewed face to face for later Waves.

## 3.7 Recommendations

In this Chapter sampling and fieldwork strategies that would be recommended, should a second cohort take place, have been explored. The outcomes of this are summarised below:

### Recommended sampling strategy:

#### Maintained schools

- A two stage sample of maintained pupils through the Annual Schools Census should take place.
- Schools should be used as the primary sampling unit (PSU), though other possibilities would include pupils' postcode sector or super-output area. Following the first LSYPE cohort, approximately 30 pupils per PSU should be chosen.
- If schools are chosen as PSUs, then small schools could be excluded from the sample frame before sampling. A (more costly) alternative would be to sample small schools separately, but with a lower selection probability. NatCen would recommend the latter option if adequate resources are available. However, small schools provide a different environment for pupils and it is plausible that pupil outcomes could be different. Excluding them, or under-representing them, would mean that the full diversity of schools would not be represented in the sample, so the decision on exclusion would need careful consideration at the design stage.
- PSUs should be stratified before sampling. The stratifiers should be chosen in consultation with the DfE, but a list of potential stratifiers to consider has been identified, including: school type, Government Office Region, urban/rural, and academic performance.
- Pupils could be stratified by variables such as ethnicity, gender, FSM, IDACI or KS3 achievement.
- Pupils from subgroups of interest should have an increased selection probability. These should include pupils from the main minority ethnic groups, but could also include, for example, those from disadvantaged backgrounds or lower achievers.
- Sample sizes should not be chosen to achieve a large Wave 1 sample, but be chosen to achieve an adequate effective sample size for each important subgroup in later Waves e.g. Wave 4. The subgroups of interest (particularly minority ethnic groups, many of which have only small proportions of pupils) and the level of precision required should first be specified by the DfE.

## Independent schools

- A two stage sample of independent pupils using the school as the PSU is recommended, and following the first LSYPE cohort, approximately 30 pupils per PSU should be chosen.
- PSUs should be stratified before sampling. The stratifiers should be chosen in consultation with the DfE, but a list of recommended stratifiers has been identified including: whether the school is single or mixed sex, Government Office Region and academic performance.
- For independent school pupils, there are very few stratification variables available at pupil-level, and NatCen therefore does not recommend stratification at pupil-level in the small number of independent schools included.

## Recommended age of cohort and fieldwork year

- NatCen recommends that the new cohort should begin when pupils are in Year 9 to maintain comparability with the first LSYPE cohort.
- To ensure that changes to policy are robustly incorporated, NatCen recommends that the study should not begin before 2012 in order to survey one of the first cohorts likely to experience the full impact of remaining in education and training until 18.
- The timing of interviews should remain the same, with pupils being interviewed in the spring and summer. However, care should be taken to ensure that fieldwork does not coincide with exam periods.

## 4 Response rates

This Chapter gives a detailed analysis of response patterns over all six completed Waves of the first LSYPE cohort and looks at how these rates have changed over time. Non-responders are compared with responders using socio-demographic variables from the previous Wave and both bivariate and multivariate analysis. By examining predictors for non-response from the first LSYPE cohort, recommendations can be made about how to minimise non-response in a second cohort. Recommendations are made alongside information from LSYPE respondents on their motivations for taking part or refusing to take part.

### 4.1 Explaining response rates

The response rates presented here are defined for individuals. The LSYPE Waves 1 to 4 data consisted of five modules: The young person (sample member) interview, household information, the main parent interview, the second parent interview, and the child history module.

Parental consent was required for the young person to take part in Waves 1 to 4. For Waves 5 and 6 this was no longer required as all respondents were aged over 16 and only the young person was interviewed. An interview was coded as 'partial' if either the main parent interview or the young person interview were not completed. If the partial interview was due to the young person not completing their section, then it was recorded as a non-response for the individual response rate.

The response rates have been calculated using a recently developed standard approach for presenting response to longitudinal studies. This framework draws heavily on the work of Lynn (2005). The following response rates are presented in this Chapter:

- Cross-sectional unconditional and conditional response rates.
- Longitudinal unconditional and conditional response rates.

For a definition of these different response rates and the calculations of each, please refer to Appendix G.

The focus of this Chapter is on study response rates rather than field response rates. Study response rates provide a better overall impression of the representativeness of the respondents that remain in the study, and can be easily compared with other longitudinal studies. See Appendix G for a definition of study and field response rates. Response rates are based on eligible young people. Also, the boost sample taken at Wave 4 is

excluded from the analysis in this Chapter because it was taken separately from the main sample and therefore has a separate response rate.

## 4.2 LSYPE study response rates

### Cross-sectional response rates

Cross-sectional unconditional rates indicate what proportion of eligible sample members in Wave *t* successfully responded in Wave *t*. Cross-sectional conditional rates are narrower as they focus on the subset of eligible sample units who have successfully responded at one or more previous Waves. Table 4.1 presents these four response rates for each Wave. Appendix G holds the definitions of these rates and sets out the calculations for each of the response rates for each Wave.

Wave	Unconditional including schools that didn't cooperate %	Unconditional excluding schools that didn't Cooperate %	Conditional on responding at Wave 1 %	Conditional on responding at the previous Wave %
1	53	73	-	-
2	45	63	85	85
3	42	58	79	92
4	39	53	73	91
5	35	49	66	90
6	33	46	62	91

- 85% of those who responded in the first Wave remained at the second Wave (that is, non-response reduced the pool of respondents by 15%).
- By the sixth Wave, 62% of those who had been responders at Wave 1 were interviewed (of those who did not become ineligible since Wave 1 due to terminating events such as moving abroad or deaths that take individuals out of the target population). This is only just lower than the National Child Development study where 63% of those who responded at sweep zero of this study responded at sweep 5 (Hawkes & Plewis 2006).
- The unconditional response rate excluding schools that did not cooperate at Wave 1 shows that by Wave 6, 46% of all issued sample members at Wave 1 remained productive (again excluding those who became ineligible since Wave 1).

- Due to school non-cooperation and individual non-response, only 53% of the eligible selected Wave 1 sample was interviewed at Wave 1. This meant that an estimated 13,894 eligible pupils were not interviewed at Wave 1 either because they were in a non-cooperating school, or they were in a cooperating school but did not respond. 59% of these 13,894 pupils were in non-cooperating schools.
- If the sampling strategy for a second cohort were to be changed by using the Annual Schools Census as the sampling frame (as recommended in Chapter Three), the unconditional response rates including schools that did not respond would be void. That is, pupils of maintained schools would be sampled directly, avoiding the problem of school non-response and hopefully increasing overall response rates (sampling issues are explored further in the next Chapter).

## Longitudinal response rates

Longitudinal response rates show response up to and including Wave 6 in relation to the original sample (minus terminating events such as deaths or moves abroad that take individuals out of the target population). The unconditional rates indicate the proportion of eligible sample members that successfully gave an interview in every Wave, up to Wave 6. The rate conditional on having responded at Wave 1 can be used to track over time how the longitudinal sample of initial Wave 1 respondents (15,431) was maintained.

**Table 4.2** Longitudinal response rates

Wave	Unconditional including schools that didn't cooperate %	Unconditional excluding schools that didn't Cooperate %	Conditional on responding at Wave 1 %	Conditional on responding at the previous Wave %
1	53	73	-	-
2	45	62	85	85
3	41	57	78	91
4	38	52	72	90
5	34	47	64	88
6	31	43	59	88

43% of those young people *issued* (and eligible) in Wave 1 responded in all six Waves, while 59% of those young people that were *interviewed* in the first Wave responded in all six Waves. This is higher than for the National Child Development Study – of those interviewed in the first Wave of this survey, only 46% responded in all six Waves.

The longitudinal response rates do not differ noticeably from the cross-sectional response rates as households that were non-responders (i.e. the main parent failed to give consent or the household was non-contactable) in a previous Wave were not reissued for the subsequent Wave. That is, Wave non-response (the temporary loss of sample members) was low and the decrease in response rates was mainly due to attrition (the permanent loss of sample members)<sup>13</sup>. Wave 6 suffered from the greatest amount of Wave non-response as it was decided to reissue non-responding young people from Wave 5 (see Table 4.2). The current thinking is that refusals from a previous Wave should be reissued at subsequent Waves. Results from the British Household Panel Study have shown that reissuing can reduce attrition from the sample over the longer term because many converted refusals remain in the sample for several waves. Furthermore, converted refusals are different from other respondents suggesting that their inclusion can improve sample composition (Burton, Laurie & Lynn, 2004). The reissuing of refusals from the previous Wave will also increase cross-sectional response rates for a second cohort. These advantages need to be weighed up against the extra cost involved in collecting the data from converted refusals, the increased complexity of weighting the data and the extent to which these converted refusals with missing data are included in longitudinal analyses.

## How these response rates have changed over time

Of those who initially responded at Wave 1, drop outs remain even between Waves. About 7% of these initial responders drop out between each Wave.

As expected, the highest category of non-responders in all five Waves (of those who responded at the previous Wave) was the refusals or opt-outs. The second highest category in all Waves (except Wave 5) was the mover category. The 'other unproductive' category was much higher in Waves 5 and 6. In Wave 5 this was mainly due to young people saying they had completed either the web or CATI interview when they had not, or having an incomplete web or CATI interview. In Wave 6 there were also a few with incomplete web or CATI interviews, but also quite a few young people away at the time of fieldwork, and more with outcome code 'other unproductive'.

Black young people were the least likely to respond in all six Waves – 43% of black young people responded in all six Waves compared with 61% of white young people.

Households in London were the least likely to respond in all six Waves. 52% of households in London responded in all six Waves, while the East of England and the South West had the highest proportion of households that responded in all six Waves at 63%.

---

<sup>13</sup> There was a small amount of Wave non-response due to young people in responding households not completing their interview but subsequently responding at some future Wave, or a non-responding household subsequently contacting the fieldwork agency to take part in a future Wave, but this was minimal.

### 4.3 A comparison of respondents and non-respondents

In this section the socio-demographic profiles of respondents and non-respondents are compared. For each Wave, current respondents are compared with non-responders (who had responded at the previous Wave).

For this section, productive households are defined as those households where both the main parent and the young person completed their interviews (for Waves 1 to 4 – for Waves 5 and 6 productive households are where the young person responded).

Because of the change of ‘respondent’ in Wave 5 (‘respondent’ here refers to the person who decides whether to refuse or cooperate – in Wave 5 the young person is approached directly instead of needing the cooperation of the main parent first), different variables were examined for the purposes of the review in Waves 5 and 6 than those in Waves 2 to 4 (along with some common household level variables). For example, in Waves 2 to 4 the main parent’s sex, age and ethnicity were compared between respondents and non-responders. In Waves 5 and 6 the young person’s sex, ethnicity, religion and whether the young person has a disability or long term illness were compared.

The characteristics of non-responders that were the same at each wave were:

The factors affecting response that were the same for all Waves were:

- A lower percentage of the non-responding households (compared with the responding households) were owner-occupiers. The lowest proportion of owner-occupiers in each Wave was in the non-contact/movers group.
- A higher percentage of the non-responding households did not have access to a motor vehicle. This was particularly so for the non-contact/movers group.
- More non-contact/movers and other unproductives were lone parent households. Also a higher percentage of young people who did not complete their interview in Waves 2-4 or who refused in Waves 5 and 6 were lone parent households compared to those who were productive.
- More responding households had the main parent in employment. In all Waves except Wave 6 the lowest percentage of households with the main parent in employment was in the movers/non-contacts group.
- Responding households had a higher percentage in managerial and professional occupations.
- Non-responding households had a lower percentage of higher educated main parents and a higher percentage of main parents with no qualifications.
- Responding households had higher income.

- Responding households were less deprived than non-responding households. The non-contact/mover and other unproductive groups had the highest proportions of households in the ‘most deprived’ quintile (except in Wave 2).
- A higher proportion of households where the young person did not complete an interview in Waves 2-4, and amongst non-responding young people in Waves 5 and 6 (especially amongst young people that refused) were households with a male young person.

In Waves 2 to 4 a lower proportion of productive households had a non-white main parent than non-responding households. Also in Waves 2 to 4, the average age of the main parent was significantly different between groups, with main parents in productive households on average being older than main parents in non-responding households.

In Waves 5 and 6, the religion of the young person differed between responding and non-responding individuals. Responding individuals were more likely to be Christians and less likely to have no religion than non-responding individuals. On the other hand, the proportion of disabled or long term ill young people was not significantly different between the groups.

Sex of the main parent was only significantly different between the groups at Wave 2 (with households that refused the most likely to contain a male main parent).

Whether the sampled young person was from an independent school in Wave 1 does not appear to have an effect on response in Waves 2 and 3 but does in later Waves. A lower percentage of non-responders were sampled from independent schools than productive individuals.

It therefore appears that in general, non-responders are likely to be of lower socio-economic status, from lone parent and/or workless households, male and from families of mixed or non-white ethnicity. This helps to identify groups at particular risk of being under-represented in future studies, particularly in a second LSYPE cohort, and this issue will be explored in further detail below. If certain groups are under-represented in the study, this can reduce the precision of findings even if such under-representation is corrected by weighting. Reissuing non-responders at the following Wave of the study could also be effective in drawing in these groups.

#### **4.4 Predictors of non-response**

In this section the socio-demographic characteristics that remain as predictors of non-response, once other factors have been taken into account, are examined using logistic regression analysis. When estimating these models two ‘hard to reach’ groups were

distinguished between – those who were difficult to contact (non-contacts) and those who were contacted, but were less likely to take part in the survey (refusers)<sup>14</sup>.

As well as the socio-demographic variables analysed in the previous section, three additional variables were also included as possible explanatory variables of non-response:

- Item non-response in the previous Wave;
- Whether the young person accepted the self completion questionnaire, refused it, or the interviewer administered it (for the Waves 5 and 6 regression analysis); and,
- Wave 5 mode of interview (for the Wave 6 regression analysis).

Item non-response is the term used to describe missing information from any one data item or question. Here the question on household income is used to measure item non-response as there is an expectation that questions about finances will suffer from high levels of item non-response because respondents tend to view such information as highly sensitive and potentially open to misuse if made public (Juster & Smith, 1997).

Appendix G shows the tables of results from the logistic regression analyses. Two regressions were run for each Wave – one for refusers and one for non-contacts/movers.

In Waves 2 to 4:

- Parents with educational levels below higher education were more likely to refuse consent on behalf of the young people than parents with a higher education level qualification.
- Lower income households were more likely to refuse than high income households.
- Those who did not answer the income question at the previous Wave were more likely to refuse at the next Wave.
- Male main parents were more likely to refuse than female main parents.
- Main parents who were economically inactive were more likely to refuse than those main parents who were employed.
- In Wave 3 households in London were more likely to refuse than households in most other Government Office regions.
- In Wave 4 households in rural areas were more likely to refuse than households in urban areas.

---

<sup>14</sup> It is well documented that the characteristics of these two groups are likely to be different (Thomas et al 2001, Nicoletti & Peracchi 2002, Hawkes & Plewis 2006).

- Households that were owner-occupied were *less* likely to be a non-contact or mover at the next Wave however further investigation found that tenure is a significant predictor of movers but not a predictor of non-contact.

In Waves 5 and 6:

- Males were more likely to refuse than females.
- Young people in households classified as lower supervisory and technical occupations were more likely to refuse than young people in managerial and professional occupation households.
- Young people from households with fewer dependent children were also more likely to refuse (this mainly arises from households with no dependent children, i.e. where the responding young person had no younger siblings).
- In Wave 5, young people were more likely to refuse if they had not accepted the self completion questionnaire in Wave 4.
- Those who had completed the interview in Wave 5 by face to face interview were more likely than those who had a web interview to refuse in Wave 6. This is probably picking up the sequential design whereby those interviewed by face to face interview in Wave 5 were (for the most part) non-responders to the web and telephone phases. Therefore this would imply that the likelihood of refusing at Wave 6 is related less to the mode of interview at Wave 5 than to the likelihood of responding at an early stage in Wave 5 – a similar finding to the self completion questionnaire in Wave 4 (see previous bullet point).
- In Wave 5 those young people living in more deprived areas were more likely to refuse than those in the least deprived areas.

Across all Waves, the younger parents were, the more likely they were to have moved or be non-contactable. Further investigation showed that this was a significant predictor of moving but not of non-contact. In Waves 2 and 4 lone parent households were more likely to be movers or non-contactable, and this predictor remained significant in both regressions for movers and non-contacts separately. Income was a significant predictor of non-contact or moving in Waves 3 and 4, with those households with lower income more likely to be non-contactable or to have moved. Similarly, the Index of Multiple Deprivation was a significant predictor in Waves 2 and 3, with households in more deprived areas being more likely to be movers or non-contactable at the next Wave.

In Waves 5 and 6, accommodation was still a predictor of non-contact or moving at Wave 5 but not at Wave 6, perhaps reflecting the fact that more young people had left home by this stage. Those young people from lone parent families were more likely to be non-contactable or to have moved in Waves 5 and 6. This is similar to the finding by Hawkes & Plewis (2006) for the National Child Development Study.

Also in Wave 6, those young people in households where the main parent had no qualifications were more likely to be a mover or non-contactable. In Wave 5, income, the Index of Multiple Deprivation and Government Office Region were all significant predictors of moving/non-contact, with those in more deprived areas or in households with less income more likely to be a mover or non-contactable.

#### 4.5 Motivation for participation or refusal – views from respondents

An additional perspective for understanding response rates can be explored via the interviews with 12 current cohort respondents that took place to inform this review. This work complements the above analyses showing the characteristics of non-responders, and provides a more contextual picture showing young people's own opinions on why they did not take part.

In this section the motivations that young people had for taking part in the last Wave of the LSYPE they participated in, the motivation for taking part initially in the LSYPE, and for those who had now refused to take part, their reason for refusal, are outlined (for summary tables of these responses see Appendix G).

##### Motivation for initial participation

Young people were asked about the reasons they first took part in LSYPE. This was six or more years ago and many were unable to recall in detail what motivated them to take part initially, but a number of reasons were stated. Interestingly a different set of factors was identified that shaped respondents' initial decisions to take part than were present in influencing participation last time (see below), and these could be seen as more coercive. They centred on perceptions that there was no choice (that participation was mandatory) or obligation (that one ought to take part).

*'Someone told me they wanted to interview me and it just kept happening each year.'* Female, BME, Short term refuser

Reasons for participation were found to change over time. This was particularly the case in relation to the role of the incentive for respondents. At the outset, when respondents were younger and money was more 'valuable' it had greater attraction than later, as they got older and gained more access to their own money (through employment, for example). The financial value of the incentive (£5 initially, rising to £8 and then £10 at Wave 7) was also a factor: £5 to a 13 year old was felt to be a reasonable sum, whereas £5 or £8 to a 16 year old had less attraction. The incentive was for some the 'hook' that brought them into the survey initially: thereafter its role became less important as they got 'the habit' or began to see participation as being important for the 'greater good'.

The roles of parents and/or friends and family in the decision to take part initially were discussed. There was some evidence that parents had more of an influence over the

decision to participate initially than later, reflecting the age of the young person at the time LSYPE first started (13/14 years old) compared with their age at the most recent Wave (18/19 years old). However, such evidence was limited by difficulties of recall.

### Motivation for participation – previous Wave

Young people were asked to think about the last time they participated in the first LSYPE cohort and why they participated. It was recognised that for some young people this was several years ago (i.e. the longer term refuser group). Two main factors shaped participating last time: those features of LSYPE that encouraged participation and reasons not affected by the survey's design. Reward and the role of the incentive in the decision making process about whether to take part in LSYPE was found to particularly important for short term refusers. In these cases, a lack of money and the opportunity to earn money by taking part was the key motivator for participation last time:

*'Well [the incentive] does make you want to do it, I suppose. Like no one does anything for free nowadays: it does encourage young people to do it.'*

Female, British Minority Ethnic (BME), Short term refuser

*'I had no money.'*

Female, White, Short term refuser

Among those who remained involved in the first cohort (i.e. current respondents) the incentive was not felt to have been the key factor. Indeed for some they had not realised they would get one. An important encouraging factor cited repeatedly among respondents was the sense that they were doing something for the 'greater good': because they thought their participation was important and would help improve things for future young people. However, there were young people whose reasons for participation were apathetic or ambivalent: there was an opportunity to take part and they thought 'why not'. There was nothing about the design of LSYPE, its purpose or implementation that motivated them to take part particularly. Such ambivalence was found across all three types of respondent but was particularly notable among longer term refusers.

*'I may as well...I just did not see any reason not to do it.'*

Male, White, Current respondent

*'You guys just sent me a letter and I just thought why not?'*

Male, BME, Longer term refuser

There was no evidence to indicate that parents or friends had influenced their decision to take part last time they were interviewed. When they had discussed it with their parents, the discussion had focused on whether the young person had been happy to participate.

This might be expected given that at the previous Wave the young people were 18/19 years old and were likely to have been making most decisions independently of their parents.

## Motivation for refusing

Short and longer term refusers were asked why they dropped out of the survey and circumstantial reasons predominated. They were busy with studying, working and other activities. Whereas initially the reasons for participation had been based on ambivalence or personal gain, as time went on the motivation to take part diminished: 'why not?' became 'why should I?' and the financial incentive diminished in appeal.

*'I suppose I just couldn't be bothered... Which is kind of bad.'*

Female, White, Short term refuser

However, there were structural features of LSYPE that contributed towards drop out. A short term refuser had become disillusioned with LSYPE because she did not receive feedback about how the results were used and therefore how her contribution was of value.

*'I can't see any point of it. I don't see any results from these interviews. I'm told each year it goes on [you're] part of something but you don't, or at least to my knowledge you don't, see any information about [what] the actual interviews [are for].'*

Female, BME, Short term refuser

A longer term refuser dropped out because he felt the questions were repetitive, covering the same ground year after year, suggesting a lack of knowledge about the longitudinal nature of the survey and why the data were being collected.

*'I felt like I'd done it before. I kept doing the same thing over and over. I kept repeating myself.'*

Male, BME, Longer term refuser

Refusals tended to take place when the interviewer had telephoned to make an appointment. However, in one case the respondent (a male, white, longer term refuser) indicated that he had not withdrawn his cooperation; rather he had not been contacted again. Had he been contacted, he said he would have continued to take part. His family situation was chaotic at the time and he now lives with other relatives. It is possible (though this was not explored in the interview) that his parent(s) refused on his behalf, without his consent or knowledge.

In terms of confidentiality, refusers (short and longer term) understood that their data were kept confidential (i.e. no one outside the research team had access to their answers linked to their names and addresses) and no concerns were expressed about data confidentiality. However, it is possible that publicised breaches of data confidentiality could result in a new cohort being less trusting, and this could affect the response rate.

## 4.6 Recommendations

### Analysis of first LSYPE cohort response rates

- Reissuing non-responders in subsequent Waves could improve response rates in a second cohort. Burton, Laurie & Lynn (2004) found the refusal conversion procedures on the BHPS to be effective in minimising attrition from the sample over the longer term. Reissues would, however, need to be appropriately weighted which can be complex and costly. NatCen recommends that non-responders in one Wave should be issued at subsequent Waves (at least for those who are not very hard refusers<sup>15</sup>). The response rate among this group is not likely to be large, so issuing them in the next Wave will not add a great deal to that Wave of the survey, but the impact of dropping refusers is cumulative in a longitudinal survey and it is now considered good practice to go back to previous Wave non-responders (Burton et al 2004).
- NatCen therefore recommends that should a second cohort take place, hard to reach groups (such as those from lone parent and/or workless households) should be specifically targeted to reduce the likelihood of bias and loss of precision in later analyses of the data. For example, since it is known that male young people are more likely to refuse, targeted and tailored intervention (such as offering incentives or tailoring advance letters) could be used on this group to ensure that they are adequately represented in the sample. Another example could be targeting households which fall in areas classified in the more deprived quintiles of the Index of Multiple Deprivation, since this information is known from the outset of the survey. Efforts to convert refusers or methods on the doorstep to evade refusal may provide the best value for money since the refusal group was the largest group of non-responders.

### Experience of respondents

- Information from LSYPE respondents indicated that when they were younger (aged 13/14) they were more influenced to take part by parents and friends and by the provision of incentives than when they grew older, who began to also consider the value of taking part as a form of altruism. There was however a high

---

<sup>15</sup> Those who refuse under any circumstances to respond and/or who state that they will never respond in the future.

degree of apathy. The recommendation is to target groups that have the lowest response rates. Young people do not actively want to drop out, but have little direct motivation to take part either. These findings highlight the importance of providing incentives and information on how the study may be used in the future to promote altruistic motivations (please also see Chapter Six and Chapter Nine).

## 5 Mixed modes

This Chapter provides an overview of the different ways of using mixed modes in a longitudinal study. It considers the impact of using mixed modes on response rates and data quality and concludes with the views of young people on their preferred mode of interview and a consideration of the optimal approach that a second cohort could adopt.

### 5.1 Strengths and weaknesses of the first LSYPE cohort mixed mode design

The first LSYPE cohort was initially designed as a face to face survey but moved to a sequential mixed modes design at Wave 5. From Wave 5, all longitudinal members were sent a letter (and email if available) with a link to the web questionnaire; those who did not complete the web questionnaire were first given the option of completing a telephone interview and then a face to face interview.

The main reason for switching to a sequential mixed modes design was to reduce data collection costs. Starting with the cheapest mode and then offering progressively more expensive modes produced substantial cost savings. This approach also offered benefits over other cost saving options in terms of frequency of data collection. For example, an alternative option would have been to maintain a wholly face to face survey but to reduce the frequency of data collection, which would have reduced the usability of the longitudinal data. It could have also reduced data quality as it then has to be collected retrospectively for a two year period and would mean that data users have to wait two years rather than one year for the next data to be released. However, inherently mixed mode designs carry a risk of differences in measurement across the modes.

Response rates reduced at Waves 5 and 6, but this would have been expected in any case given the changes taking place in most longitudinal members' lives at the time e.g. respondents aged 17-19, moving out of the parental home, starting higher education, going into employment.

### 5.2 What are the alternative mixed mode designs?

Increasing pressures of rising costs of survey operations and falling response rates have led many studies to explore the potential benefits of combining different modes of survey data collection. While the first LSYPE cohort adopted a particular mixed mode design, it is not the only option available and this section outlines different mixed mode designs that could be considered for a second cohort.

Two possible mixed mode designs are available for collecting the same data from different sample members using different modes in a single Wave of the survey. These

different modes can either be offered upfront (a concurrent design) where respondents are given the choice of mode, or in a predefined sequence, so that all sample members are asked to complete the survey using one mode before moving on to another mode for non-responders (a sequential design). These designs are proposed in an attempt to reduce costs and/or increase response. Depending on how the modes are mixed, it may also be possible to increase response rates and reduce non-response bias.

### **Concurrent design: respondents given a choice of mode upfront**

Survey estimates can be biased if non-respondents are different from respondents in relation to what the study is trying to measure. This non-response error tends to vary across modes to the extent that different sample members are more likely to respond to a given mode (for example, more web-literate people are more likely to respond to a web questionnaire). If different sample members are more likely to respond to one mode than another, then it should be possible to use mixed modes to maximise response and to increase the representativeness of the responding sample. This has led many to explore the possibility of providing respondents with a choice of data collection modes. Indeed, respondents reported in interviews about the first LSYPE cohort that a choice of modes was important and they liked having a flexible approach.

However, there is no evidence to suggest that this design actually improves response rates or reduces the risk of non-response bias. Various experimental studies have shown a lack of evidence that a choice of mode will increase overall response (Dillman et al, 1995; Lozar Manfreda et al, 2001). One possible explanation for this is the 'responding through the mode at hand' principle; e.g. if sample members are sent a postal questionnaire with the option of completing the questionnaire online, they are more likely to complete the postal questionnaire than to make the extra effort to go online (Holmberg et al, 2008). Although concurrent designs do not increase overall response, it has been suggested that offering a choice of mode could possibly build goodwill and improve attitudes toward survey participation, which could be beneficial for a longitudinal survey (de Leeuw et al, 2008).

### **Sequential design: respondents offered different modes in predefined sequence**

A sequential mixed mode design can be used to minimise costs without sacrificing response by using the cheapest data collection mode first, before proceeding to the use of increasingly more expensive data collection modes among the remaining non-responders (Hochstim, 1967; Japac, 1995; Voogt & Saris, 2005).

### **Switching mode between Waves of a longitudinal study**

Longitudinal surveys provide the opportunity to use different data collection modes at different time points. The main reasons for using different data collection modes at

different time points could be practical issues or costs. For example, face to face interviews may be preferred at the outset of a longitudinal study, in order to introduce the study and gain the trust of respondents, whereas in later Waves it may not be considered as important to have personal contact with a respondent who has already participated. As a second example, it may not be feasible to collect data by telephone or email at the first Wave because the relevant contact information is missing or inadequate. However, this contact information could be collected from the respondent during a face to face interview, thus providing the opportunity to collect the data using other, cheaper modes at subsequent Waves. Depending on what data items are to be collected at different time points, one mode may be more suitable than another (e.g. a short list of simple factual questions could be collected by post).

There can be serious consequences of switching modes between Waves when surveying the same sample over time however. For example, if multiple modes are used for different Waves of a longitudinal survey (e.g. first face to face interviews and then telephone interviews), it is almost impossible to decide whether changes in respondents' answers are real changes or results of a change in mode: '*Time effects and mode effects in the results are fully confounded*' (de Leeuw, 2005).

### **5.3 Characteristics of different modes**

Having considered concurrent, sequential and the effect of switching modes between Waves, in this section the strengths and weaknesses of different modes themselves (i.e. face to face, telephone, postal and internet survey) are considered (summarised in Table 5.1).

**Table 5.1 Summary of the strengths and weaknesses of different modes of data collection**

Design Parameter	Face to face	Telephone	Postal	Web
<b>Cost of data collection</b>	Usually most expensive method.	Usually around 50-70% of face to face cost for same interview.	Relatively cheap (but questionnaires need to be kept short and simple).	Cheap (no print, interviewer or data input costs).
<b>Amount and type of resources required</b>	Specialised fieldworker skills and field-force management resources needed.	Specialised interviewer skills and management resources needed.	Operational resources for managing mail-outs returned questionnaires.	Programming and web hosting resources needed.
<b>Timetable considerations</b>	May require several months unless respondents are easily accessible or 'captive'.	A potentially fast mode of data collection, but depends on respondent availability.	With response reminders, may require several months.	Usually the fastest mode of data collection, but likely to require postal/email reminders to achieve acceptable response.
<b>Operational control</b>	Best for control of field sampling and data collection.	Good for interviewer supervision, but respondent tolerance may be limited.	Few means of controlling how questionnaires are completed.	Question routing and ordering can be controlled by programming.
<b>Amount/complexity of data to be collected</b>	Best/mandatory for long and complicated questionnaires.	Limitations on length and data collection complexity compared with face to face.	Weaker for groups with poor literacy or motivation, but can be good for experts.	Requires computer and language literacy. Complex routing can be programmed into web questionnaires.
<b>Data Quality</b>	Best for complex topics and issues. Computer assistance improves quality. May incur interviewer effects.	Good for simple factual and attitudinal questions. Computer assistance improves quality. Interviewer effects less likely.	Worst for missing data, routing errors, misunderstandings.	May include prompts if questions are missed and data validation can be programmed into web questionnaires.
<b>Statistical efficiency</b>	To reduce fieldwork costs less efficient clustered samples needed for national surveys.	Does not require clustered samples.	Does not require clustered samples.	Does not require clustered samples, but may have sampling problems (i.e. coverage).
<b>Expected response rate</b>	Usually gets highest rate.	Likely to be 10-30% lower than face to face.	Can be well below 50%.	Limited evidence, but generally likely to be a low response rate: may be higher among computer literate and young respondents.

*Adapted from the Magenta Book: Guidance notes on policy evaluation*  
[http://www.nationalschool.gov.uk/policyhub/magenta\\_book/](http://www.nationalschool.gov.uk/policyhub/magenta_book/)

## Response rates and non-response bias

Response rates tend to vary by mode with face to face interviews having the highest response rate, followed by telephone interviews, and then postal and web questionnaires. Meta-analytic evidence suggests lower response rates for web than other modes (Manfreda et al, 2008; Shih and Fan, 2008). However, the magnitude of difference in response rate by mode varies depending on survey topic, population group, availability of relevant contact information, respondent burden, and the amount of effort that is made to maximise response within each mode (e.g. number of contact attempts, refusal conversions and incentives). One relevant example of how response rate for modes may differ by population group is that of respondent age; there is some evidence that compared with older respondents, younger people are more likely to state a preference for web over postal surveys (Miller et al, 2009). In addition to response rate, non-response bias also differs by mode; face to face surveys tend to achieve a more equal response rate across population subgroups than telephone, postal and web surveys.

These differences in response rates by mode are important factors to consider when choosing data collection modes for a longitudinal study, particularly for the first Wave of the study where the greater part of non-response occurs. A face to face interview is usually the preferred choice for the first Wave of a longitudinal study due to the benefits that the presence of an interviewer brings, such as screening in eligible respondents and gaining their cooperation with the study.

## Question design

When there is no obvious primary data collection mode for which to design questions, or when the data collection mode may be changed Wave by Wave in a longitudinal study, there are three main strategies for minimising mode effects in terms of measurement: mode-specific design, uni-mode design and generalised mode design (de Leeuw, 2008).

For the mode-specific design, questions are designed separately for each mode, making use of the unique features of that mode (e.g. use of show cards in face to face interviews or supporting diagrams in web surveys). The rationale is to reduce overall error. However, this is based on the assumption that the same concept is being measured across all modes but with different levels of accuracy (i.e. no systematic bias). This may not be the case. For example, the use of branching of response options (optimum design for telephone mode) and no branching of response options (optimum design for self completion modes) produces systematic differences in responses between modes. Therefore, this strategy for minimising mode effects comes with the burden of proof on the designer.

An alternative strategy for minimising mode effects when there is no primary data collection mode is uni-mode design: ‘the writing and presenting of questions to respondents in a way that assures receipt by respondents of a common mental stimulus, regardless of survey mode’ (Dillman, 2007; p. 232). A uni-mode design seeks to produce questions that share design features across modes. Dillman (2007) has produced a list of recommendations for uni-mode design, such as making all response options the same across modes and incorporating them into the stem of the question, reducing the number of response options, using the same descriptive labels for response categories, developing equivalent instructions for skip patterns, etc. However, there is a risk when trying to standardise the features across modes that this will result in sub-optimal design for all modes.

The final suggested strategy for minimising mode effects when there is no primary data collection mode is the generalised mode design. Questions are purposively constructed to be different in different modes with the aim of achieving cognitive equivalence of the perceived stimuli, thereby resulting in equivalent answers across modes (de Leeuw, 2008). Although there has been some work done in this area (for example, Christian et al, 2008), it is still early days and much more empirical research is needed to understand what represents the same stimulus across different modes.

Results from the in-depth interviews with the research consortium that collected the data for the first LSYPE cohort suggest that the preferred strategy for minimising mode effects was uni-modal design. However, this would only have been possible for new questions that were being introduced when the switch to mixed modes is made. Prior to Wave 5 the primary data collection mode had been face to face interviewing, and questions were designed optimally for this mode. The research consortium commented on the difficulty of using uni-modal design for mode combinations that used visual versus aural stimuli.

Development work can be undertaken to identify whether particular questions can or cannot be asked in different modes, and to design the survey, taking these findings into account. As part of this review, an Expert Panel examined a wide range of questions that have been included in the first LSYPE cohort study. A summary of the Panel’s findings can be found in Appendix I with general principles on how to use such question types in a mixed mode survey.

## **5.4 Respondents’ mode preference**

Young people who had taken part in the first LSYPE cohort were asked about their preferred mode of interview. Preferences were varied and reflected the young person’s circumstances. Offering a number of different ways in which young people could participate in LSYPE was felt to be important and young people liked this flexible approach.

When asked about their preferred mode, those who chose face to face interviewing liked it because they felt secure, they could see who they were giving their answers to, it felt personal, it was easy (they didn't have to write/type and they felt more confident in answering) and they liked the fact that they could build up a rapport with the interviewer. Those who chose the web (online) said it was because it was convenient (you could do it where you want to and when you want to), it was easy to complete, you didn't have to talk to anyone, it was private (they felt more comfortable being asked sensitive questions) and that their answers could not be overheard by others. Where telephone interviewing was preferred respondents said it was convenient as you could do it where you want (if called on mobile), it was private and they liked being able to build a rapport with the interviewer.

A face to face interview was felt to be appropriate for the first interview and it was felt that younger people would be more comfortable with this format. However, as young adults there was a view that telephone and/or web were more appropriate, affording greater flexibility (and control) over where and when the interview took place. Those with busy lifestyles tended to prefer web or telephone. Confidence was an issue for those who struggled with reading and writing (skills felt to be important in completing a web questionnaire) and felt nervous using a telephone.

Development work should therefore be undertaken prior to a survey taking place (before the pilot stage) to identify whether particular questions can (or cannot) be asked in different modes, and design the survey taking these findings into account.

Young people made a number of recommendations for improvements to LSYPE, which are listed below:

- Maintain flexibility in the way respondents can participate in the survey – face to face, telephone and web.
- Do not place too much emphasis on the web: some respondents who don't have access or do not like using it.
- Face to face is the most appropriate means of participation at Wave 1.
- Consider offering a 'dial in for an interview' service, in later Waves, whereby respondents can call a free telephone number to do an interview at a time that suits them. This option might be useful for particularly busy respondents.

## 5.5 Mixed mode approaches and the LSYPE

It is recommended that should a second cohort take place the primary data collection mode from Waves 1 to 3 should be face to face interviewing. Self completion forms (paper or CASI<sup>16</sup>) could be used during the face to face interview to collect sensitive information (as occurs in the first LSYPE cohort). Questions that are susceptible to mode effects could also be included in the self completion form if it is known that they will be repeated in later Waves using self completion modes (e.g. web).

Reasons for proposing face to face interviews until Wave 4 are the need for signed parental consent at each Wave until the age of 16; the collection of survey data from parents as well as the young people and gaining acceptance of the study from sample members at the outset. Face to face contact is considered necessary to inform parents about the benefits of the survey and to collect their signed consent forms prior to collecting information from the young person. Furthermore, it can be difficult to get multiple people in the household to complete a web questionnaire or telephone interview and a large proportion of households could well be allocated to a face to face interview anyway. As soon as one person in the household requires a face to face interview, the added cost of face to face interviewing for the other household members is minimal. For these reasons, it is recommended that should a second cohort take place a mixed mode design should only be used for those Waves where parental consent and parental participation are no longer required. However, given that it may be difficult to achieve face to face interviews with both parents, it may be necessary to collect data from parents in other modes from the outset of a second cohort.

Although non-response remains an issue of concern for the duration of the study, the drop out after the first Wave of data collection is far less severe. Relevant contact details (e.g. stable addresses, email addresses, and fixed and mobile telephone numbers) can be collected at each Wave to allow other modes to be used after Wave 3 and this information can be updated throughout the lifetime of the study.

The main advantage of using mixed modes after Wave 3 would be the potential for substantially reducing data collection costs. Cost savings could be achieved by switching from expensive face to face interviewing to a less expensive mode, such as web or telephone. However, switching to either web or telephone could have a negative impact on coverage and response which would be detrimental to the perceived integrity of the longitudinal data.

---

<sup>16</sup> CASI = Computer-Assisted Self-Interviewing

An alternative recommendation for a second cohort would be a sequential mixed mode design, which would use a less expensive mode for all longitudinal members before switching to more expensive modes among non-respondents after Wave 3. Not only would this reduce costs, but an equivalent (possibly higher) response rate to a complete face to face data collection Wave could be achieved if the sequential design is concluded with face to face interviewing. An added advantage of switching to mixed modes after Wave 3 is that this would coincide with an age (from 16) at which young people are becoming more mobile and the sample is less clustered. It would therefore be easier and cheaper to track and interview these young people using, for example, web and telephone methods.

Switching to a sequential mixed mode design after Wave 3 could however increase the risk of differences in measurement between modes. To minimise this risk, the second cohort should be designed upfront as a truly multiple mode longitudinal survey and the questionnaires designed accordingly. It would therefore be desirable to map in advance the range of questions/topics that would be asked at each Wave<sup>17</sup>. Those questions that are only to be asked in Waves 1-3 could be designed optimally for face to face interviewing (or self completion). Other questions could be designed so that they are portable across modes. How this is achieved may differ for different questions. The most obvious method would be a uni-mode design but mode-specific design and generalised mode design could also be considered.

Nonetheless, some questions could not be used across all three modes (i.e. web, telephone and face to face) without introducing differences in measurement, whereby responses to questions vary by mode of data collection due to the characteristics of different modes. Specific examples are included in the expert Panel's recommendations (Appendix I). One option would be not to ask these questions in the modes where they are problematic; but this could introduce the risk of bias when using a sequential mixed mode design, where different modes are used in a single Wave of the survey. Alternatively, if there were a number of questions which could not be asked in a particular mode without introducing the risk of measurement error in comparison with other modes, this mode could be excluded during Waves when these questions are asked. Best practice would be guided by a map of questions/topics; when these questions/topics need to be asked; and the modes in which these questions/topics could be asked without introducing measurement differences.

---

<sup>17</sup> It is recognised that some flexibility for introducing new questions and topics during the lifetime of a longitudinal study is required to optimise the relevance of the data for policy makers and other data users.

## 5.6 Recommendations

- NatCen recommends that if a second cohort takes place, the young people should be interviewed face to face for the first three Waves. This is because: it is necessary to obtain consent from parents during these Waves before an interview can take place with young people under 16; the first Wave of a survey is more successful in terms of response rate if the mode is face to face; and, to gain acceptance of the survey among sample members. However, due to possible difficulties collecting data from both parents, a mixed mode design for parental data collection may be necessary from the outset for any parent who is absent or unavailable for face to face interview. In such cases, the questionnaire would be administered in another mode, such as telephone, postal or web.
- A mixed mode approach for the young people after Wave 3 is recommended because it could result in considerable cost savings. The recommendation would be to conduct a sequential mixed mode design with the cheapest mode used first and then more expensive modes worked through for non-responders. For comparability with the first LSYPE cohort, data could continue to be collected from parents after Wave 3, also using a mixed mode design.
- A second cohort should be designed from the outset to be mixed mode, and be designed so that questions with the greatest risk of measurement error between modes are not asked in mixed mode Waves if possible. Any questions that are likely to be asked in different modes at some point during the study should be designed to be portable across modes.
- A consequence of a mixed mode approach to question design and use may be to reduce comparability with some of the data collected from the first LSYPE cohort. Development work will be able to assess how many questions this will affect, and whether or not the benefits of a mixed mode question design outweighs the disadvantages of non-comparability between the cohorts.

## 6 The use of incentives

To help make informed decisions around the use of incentives in longitudinal surveys, and particularly a second cohort of LSYPE, this Chapter presents an overview of how incentives have been used in cross-sectional and longitudinal studies, starting with how they are used in the first LSYPE cohort. Of specific interest is how incentives may affect response rates, fieldwork costs, sample composition and data quality. Finally a set of recommendations is outlined.

### 6.1 Overview of key longitudinal studies and their incentive methods

Incentives have become one of the main tools available for surveys to encourage respondent participation and have been used by the first LSYPE cohort since Wave 1. They are increasingly used as a way to maximise overall study response rates and in an attempt to reduce non-response amongst specific groups. Incentives often take the form of either a gift or money, but can also include charitable donations or lottery draws. They can be administered in advance of the survey taking place on an unconditional basis, or after the survey has been completed to show appreciation for the time given (on a conditional basis). Influential theories of social exchange, reciprocity, and intrinsic motivation were examined to inform this review to help evaluate whether the impact of incentives on adult behaviour can be generalised to that of young people (please see Appendix J). It is also important to note that – particularly in studies of young people – incentives should not be a substitute for interesting and engaging content in the study, which is also likely to affect respondents' motivation for taking part.

#### The first LSYPE cohort – monetary incentive

The first LSYPE cohort used a monetary incentive. All cohort members of the first LSYPE cohort who completed a Wave 1 interview were given a £5 high street voucher. For Waves 2 and 3, administration of the incentive changed so LSYPE cohort members were sent an unconditional £5 voucher with their advance letter at the start of each Wave. At Wave 4 the value of the unconditional high street voucher was increased to £8. From Wave 7 this was increased to £10.

Most longitudinal studies that use monetary incentives increase the amount over time to combat falling response rates. It is generally felt that the incentive needs to remain meaningful to respondents relative to the current costs of living, and reflect the level of burden in terms of interview length. It is expected that increases in incentive amount over time signify reward for long term commitment to the study.

Some examples of British studies that have used monetary incentives are the British Household Panel Study (BHPS), and the English Longitudinal Study of Ageing (ELSA).

Examples of international studies that have used monetary incentives are the Panel Study of Income Dynamics (PSID), the National Longitudinal Surveys of Youth (NLSY), the Health and Retirement Study (HRS) and the Household Income and Labour Dynamics in Australia (HILDA) survey. British studies have tended to keep the incentive rates fairly low in comparison with their international counterparts. For example, PSID currently pays respondents \$1 per interview minute, and both NLSY and HRS pay around \$50 per interview, whereas BHPS pays £10). Further details of the incentives used in these studies can be found in Appendix K.

## Studies without a monetary incentive

Although current opinion tends to advocate monetary incentives, some longitudinal studies have chosen not to adopt them. Often this is a result of budget restrictions or because response rates remain sufficiently high without them. However, other contributory factors may include worries about expectation effects (once respondents' receive an incentive they will come to expect it), or general ethical concerns. Instead, surveys of this kind rely on respondents' sense of altruism, or the possible role played by longitudinal identity on willingness to cooperate.

Examples of studies which do not use monetary incentives include the Canadian Survey of Labour and Income Dynamics (SLID), and the Swiss Household Panel (SHP). Of more relevance to LSYPE are the British Birth Cohort Studies (NCDS, BCS70, MCS) and the National Longitudinal Survey of Children and Youth (NLSCY) which have opted to give out small age-appropriate gifts to children – including fridge magnets, pens, rulers, and wall charts.

## Use of 'additional' incentives

Some longitudinal studies covered by this review have started to build in 'additional' incentives in return for completion of specific tasks or for use at the interviewer's discretion. The list below shows some of the items offered:

- Finders fees for family members who provide a new address for sample members who have moved (PSID);
- \$10 for returning an address confirmation card mailed to longitudinal members each year (PSID);
- Gifts in kind / pens / diaries (BHPS);
- Flowers / chocolates for birthdays or special occasions (BHPS);
- Food up to a maximum of \$20 – e.g. family size pizza (NLSY);
- Offer of alternative site for interview combined with offer of a free meal (e.g. coffee house, restaurant chain, fast food outlet) (NLSY).

## 6.2 The effectiveness of different incentive strategies

To fully test the effectiveness of different incentive strategies on response rates, some studies have set up incentive experiments. Some of these relate to the incentive used at the first Wave, while others look at later Waves.

### Use of incentives at the first Wave

Some of the experimental work at the first Wave of studies has focused on whether an incentive should be offered at all, and if so, what the incentive amount should be to influence participation rates. The Understanding Society team set up an Innovation Panel to provide a platform for methodological testing which could inform the design of other studies as well as US itself. At the first Wave, an unconditional monetary voucher was sent with the advance letter to three experimental groups. Adults in the first group were sent a £5 voucher, the second group were sent a £10 voucher, and the third group were sent a £5 voucher which increased to £10 if all adults in the household participated. The £10 conditional incentive response rate was marginally higher than the £5 one. The study team therefore chose to send a £10 voucher with the advance mailing to all those eligible at its first Wave.

In 1996, the Study of Income and Program Participation experimented with the amount of incentive offered at its first Wave (James, 1997). They concluded that the amount paid has an independent effect on response, but that it also needs to reach a certain value level before influencing the decision to participate. Once this happens, the effect of the incentive at the initial interview is then upheld over the first three Waves.

### Use of incentives at subsequent Waves – but not at first Wave

Some longitudinal studies have experimented with incentives at later Waves to boost response rates. The British Election Study (BES) and the 10<sup>th</sup> cohort of the England and Wales Youth Cohort Study (YCS) did not initially offer an incentive in the first Wave (referred to in YCS as a sweep) and results showed that there is no carry-over effect from one Wave to the next on likelihood to respond.

This experimenting showed that the unconditional incentives reduced later attrition, and item non-response in comparison to the conditional incentive. Also the use of incentives did not impact on attrition bias across a range of respondent characteristics (i.e. composition of sample was unaffected by the different treatment groups). Item non-response was reduced more in postal than telephone mode.

Jackle & Lynne (2007) found that changes in the treatment from conditional to unconditional and from telephone to postal did not affect the outcomes in later Waves. Therefore it is possible to alternate or change the experiences of respondents at a given Wave without it impacting on willingness to participate at later Waves. Despite less

committed panel members dropping out, they found little evidence that the remaining sample became less sensitive to the effect of an incentive over time.

## Increasing incentive after Wave 1

Some studies have included an incentive at the first Wave and then experimented with increasing its value at later Waves for either the whole sample or specific subgroups. For example, the Health and Retirement Study and the British Household Longitudinal Study both ran experiments to test the effect of increasing the incentive value from an already established level. Details are found in Appendix K.

Laurie & Lynn (2008) concluded that the incentive does not need to be increased by a huge amount to have positive effects on response. In this case, sample members had become accustomed to the incentive amount, so increasing it slightly had beneficial psychological effects independent of the value of the increase.

The marginal increase in incentive value also had the greatest effect on specific groups – namely prior Wave non-responders, and those in the youngest age group. Arguably it encouraged people to take part who would not have otherwise, and in persuading those who were likely telephone interviewees to be interviewed face to face. Although, a face to face interview is more expensive than a telephone interview it is certainly preferable in this circumstance in terms of data quality and achieving a better sample.

## Targeting specific groups

It is clear from the available literature that particular groups appear more sensitive to the effect of incentives – e.g. lower income, lower educated (James, 1997, Stratford et al, 2003). These groups are also commonly found to be more likely to non-respond. As a result, some studies like the Health and Retirement Study and the Survey of Income and Program Participation have chosen to target non-responders with higher incentives to encourage participation and help reduce sample bias. In addition, the National Longitudinal Study of Youth has shown that differential incentives can be used in conjunction with mode to promote greater survey cost effectiveness (See Appendix K for a full description).

During the depth interviews with respondents it transpired that incentives were an important ‘hook’ for gaining initial participation and for encouraging continued participation and that vouchers were generally preferred to cash or cheques, though one respondent expressed a preference for cash. Offering a choice of where vouchers could be spent was felt to be important (i.e. High Street, cinema, iTunes). Not all respondents had (easy) access to the internet however which should be noted when choosing incentives.

## 6.3 Recommendations

There is overwhelming evidence from numerous experiments carried out over the last 30 years that respondent incentives increase response rates, especially amongst specific groups. Younger people have been found to be particularly sensitive to their use, as are low income and lower educated groups. The ethical implications of such incentives must be carefully considered, given targeted incentives could alienate young people who do not fall into these latter two groups.

There is strong evidence to suggest that incentives can be cost effective (Singer, van Hoewyk and Couper, 1998; Lynn, Thomson and Brook, 1998; Salathiel and Nicolaas, 2004). Incentives can reduce the total number of calls made at first issue to achieve the interview as well as the total number of cases that need to be reissued. Prepaid respondent incentives are especially important in panel surveys because of the critical need to recruit a high proportion of the eligible population into the initial round of measurement. Respondents interviewed as part of this review concurred that incentives are an important motivator for taking part. NatCen therefore recommends that incentives are offered if a second cohort occurs.

### Types of incentives to offer

- Cross-sectional experimental work suggests that money is more effective at increasing response rates than a gift. However, no longitudinal studies were found that had experimented with different types of incentive - most had focussed on changing the value, or the conditions upon which monetary incentives are given (e.g. from conditional to unconditional). The type of incentive used also may be affected by age. Money or vouchers have been found to be effective with teenagers attending out of school time programs, and has been used with other longitudinal studies of young people (e.g. first LSYPE cohort and Youth Cohort Study). NatCen therefore recommends that a monetary incentive is offered.

### Value of incentive

- The first LSYPE cohort members received a £5 incentive in the first wave, increasing to £8 in Wave 4 and £10 in Wave 7. In the case of a potential second cohort NatCen recommends starting with a higher incentive at the first Wave (e.g. £10) to attract people to the study and this amount should remain at the same level for the first 3 Waves.

## Incentives at subsequent Waves

- Evidence from the BHPS showed that if the same incentive value is maintained for many Waves, marginal increases at a later Wave can instigate an up-turn in response rate. Conversely, large monetary boosts (often targeted at non-responders) can also be effective with little evidence of a carry-over effect to subsequent Waves. Experimental evidence also suggests that the administration of the incentive and mode of interview can be changed at later Waves without having a detrimental effect on response rate. Therefore, NatCen suggests a marginal increase from £10 to £15 from Wave 4, as the study switches to mixed mode.

## Unconditional or conditional incentives?

- Cross-sectional evidence suggests that unconditional incentives are more effective at increasing response rate than conditional incentives. Two meta-analyses have also collated a wide-range of experimental studies to show that prepaid (unconditional) incentives are more effective than promised (conditional) incentives in increasing response (Singer et al, 1999; Church, 1993). Although, most of the longitudinal studies have opted for conditional incentives, NatCen would recommend unconditional incentives at all Waves, particularly at Wave 1 because of the critical need to recruit a high proportion of the eligible population into the initial round of measurement.

# 7 Interviewing parents

Having reviewed significant aspects of the first LSYPE cohort and made recommendations for the design of a second cohort, a number of additional considerations remain. This Chapter focuses on the issue of interviewing parents. It includes a discussion of which parent to interview, provides examples of studies that have interviewed a second parent, outlines the potential uses of parental interview data and provides recommendations for a second LSYPE cohort.

## 7.1 Who should be interviewed?

For the first four Waves of the first cohort, interviews were conducted where possible, with a young person, main parent, and second parent in each household. This section explores whether having three interviews in a longitudinal study of young people is efficient, or whether a young person and main parent interview would suffice.

### Who was interviewed at each Wave of the first LSYPE cohort

In the first LSYPE cohort it was initially planned to interview main parents (usually the mother) in Waves 1 and 2 moving to telephone interview of the main parent only at Wave 3. However, due to the success of the first two Waves extra funding was secured that enabled the project to continue interviewing the main parents face to face for Waves 3 and 4. This method helped retain high response rates, ensuring many of the respondents remained with the study from Wave 3.

Second parents were interviewed face to face in Wave 1 and then interviewed over the phone if in Wave 2 if interviewers were unable to collect data in Wave 1. For full details of who was interviewed at each Wave, definitions of main parents and how these changed, and the length of the interviews see Appendix L.

At Wave 5 it was felt that the parental interviews were no longer critical - much of the background and household information had been collected in previous Waves and a value for money case could no longer be made. In addition, young people were old enough to consent to answering questions without parental approval. Therefore parent interviews were dropped. Once LSYPE involved interviewing just a single respondent per household, mixed mode interviewing also became viable thereby saving a great deal of resource.

## 7.2 Other longitudinal studies that include a parental interview

See Appendix L for details of the following comparative longitudinal studies that include a parental interview:

- Growing Up in Scotland (GUS)
- The National Child Development Study (NCDS)
- Millennium Cohort Study (MCS)
- The 1970 British Cohort Study (BCS70)
- The Avon Longitudinal Study of Parents and Children (ALSPAC)
- Growing Up in Australia (LSAC)
- The National Longitudinal Study of Adolescent Health

### Data collection methods with parents

Various methods can be used for collecting data from parents. There are face to face interviews with both parents conducted in the home, proxy interviews with main parents to collect data on second parents or self completion questionnaires. There are further options available when more than one person in the household has to be interviewed.

The English Longitudinal Study of Ageing (ELSA), for example, adopts concurrent interviewing whereby couples are interviewed together by Computer Assisted Personal Interviewing (CAPI). The interviewer asks a block of questions to one member of the couple, and then repeats the same block of questions to the other. This has a time saving element when compared with two separate interviews, and is also made more enjoyable for couples who share the interviewing experience. It is quite feasible that concurrent interviewing, or some variation of it, could be adopted for parental interviews in a future LSYPE. While the effects of the concurrent method on data quality have yet to be fully tested, it is widely accepted that it enhances the interview process, and there is no evidence that it significantly biases the results. As with all these methods, there are disadvantages. Confidentiality and the provision of socially desirable answers are difficulties with concurrent interviewing. Similar to interviewing by proxy, the reliability of questions on topics such as income, qualification and parental style may be questionable.

Another option would be to collect information from the parents via the internet. Obviously not all parents would have access to the internet and not all those with access would complete the questionnaire online, so other modes would have to be used alongside this. For example, offering the possibility of using a web based survey for parents with internet access with the possibility of a postal interview for parents without access. Telephone interviewing could then be used as a follow up for non-responders. Depending on the response rate and the budget, face to face data collection could then

be used as a last resort. This kind of sequential mixed mode design is discussed in more detail in Chapter Five.

## What are the parent interviews used for?

Parental information from LSYPE has been used to help contextualise young people's experiences and complement the information provided by the young people themselves. It enables researchers to build up a fuller picture of the lives of the young people, and examine the impact this has on young people's own experiences. For example, socio-economic information is key to helping us understand social mobility and can be derived from data provided during parental interviews. This information has been used extensively in LSYPE analysis. Even when it is not the main focal point in the study, information from parents enables the researcher to control for family circumstances providing a more accurate picture of other influences such as individuals own or their school's characteristics.

Most recently the main parent interview has been used as one of the key outcome variables in a project commissioned by DfE looking at the characteristics of bullying victims in schools. LSYPE collects information about whether the young person reports being bullied. It also collects data from parents about whether they think their children are being bullied. This added an invaluable dimension to the study enabling the researchers to explore whether parental awareness of bullying at an earlier age made any difference as to whether the young person was bullied at age 16. The analysis showed that young people whose parents had reported them being bullied at age 14 and 15 were more than twice as likely to *not* be a victim of bullying at age 16.

Many studies that use LSYPE data have examined the influence of parents on young people's experiences, looking at their impact across a diverse range of different outcomes including, for example, their attainment, aspirations, engagement with education, alcohol consumption, and engagement in risky behaviours. For a list of other recent studies using this data see Appendix C.

### 7.3 Including a second parent interview

There is also a benefit in collecting information from second parents (usually the father), however there are trade offs between cost and depth that would need to be considered when making this decision. Whether or not second parents were interviewed would depend primarily on what data was anticipated being collected balanced with the logistics and cost of collecting this data.

For example, collecting data from both parents would be imperative for studies examining parenting styles and attitudes in a second cohort. Whilst there is existing evidence on less traditional families (i.e. lone parent and 'broken' homes) there is less evidence on the family as a unit, how young people are situated within that unit, and how

this changes over time. In addition, there is a lack of detailed information on the role of fathers in the family. Collecting data from both parents would be important to address these gaps. However there can also be difficulties in collecting data from a second parent, which is more often the father. For example, the NCDS attempted to get information from both parents, although in reality data on both parents was mainly provided by the mother, as the father was often unavailable to complete an interview.

There is therefore a strong argument for collecting data from a second parent as well as a main parent. On the other hand, second parent data is often not used in practice because of low response rates. If the study was designed to elicit a high enough response rate then the data could be highly valuable. Alternatively (depending on the information to be collected) it could be collected by proxy from the main parent, cutting down time and cost considerably. The outcome datasets would also be more straightforward as there would be no need to have a separate file for second parent interviews. However, as already noted above, this will depend on the data being collected.

Ideally, in a second cohort, if the interviewer was making a home visit to collect parental consent as well as to interview the main parent, then they should also try to interview the second parent during this visit. This would be the same procedure as that adopted in the first LSYPE cohort. If it is not possible to interview the second parent during these visits, then some other strategy would be required to boost the response. NatCen suggests that these second parents with internet access are asked to complete the web questionnaire and those without, to complete a postal questionnaire. It would be most practical for the interviewer to leave a postal questionnaire for all unproductive second parents, with a URL address and password for a web questionnaire. It might also be possible to collect an email address for the second parent, if available, and the fieldwork agency could send an email to the second parent with the URL address and the password embedded. This should increase the response to a web questionnaire. A final stage would be to contact the remaining non-responding second parents by telephone and attempt to collect the information using a telephone interview. The advantage of trying to collect the data by internet or post prior to a telephone interview is to reduce the overall costs (see Chapter Five on mixed mode designs).

## 7.4 Recommendations

With regards to the parental interview:

- NatCen recommends that interviews with both main and second parents are conducted for at least the first three Waves, but recognise this will be depend both on the desired content of the study, and the resources available.

- Which topics are of key interest to policy would have a guiding influence on whether to conduct interviews with both parents in a second cohort. However, due to the recent change in Government it is not currently clear what the priorities would be. For example, if the impact of different parenting styles were of interest then interviews with both parents would be very important.
- NatCen also recommends testing the accuracy of collecting information by proxy as there are clearly benefits in conducting a proxy interview despite the bias this has on some questions.
- There are a number of options regarding the mode of interview for second parents in a second cohort, of which three are outlined below. This will depend on the resources available and content of the interview. It is assumed that the main parent would be interviewed face to face for the three Waves in line with the young person interviews. Options for second parent interviews include:
  - Face to face interviews conducted with the main and second parents at Wave 1. Interviews in following Waves are then conducted with the main parent only (face to face).
  - Face to face interviews conducted with both the main and second parents at Wave 1. Then face to face interviews conducted with the main parent in following Waves and interviews with second parents at Waves 2 and 3 conducted via the internet or post (and by telephone for non-responders).
  - There are no interviews with the second parent however basic information is collected by proxy interviews with the main parent.
- Second parent interviews at some or all Waves could be conducted via telephone or web to reduce costs.

## 8 Data linkage

The Chapter outlines the advantages of data linkage when conducting large scale longitudinal surveys. Datasets that are of potential value for linking to a second cohort are described along with information regarding how to negotiate access. The Chapter concludes with recommendations regarding data linkage.

### 8.1 Strengths and weaknesses of first LSYPE cohort and data linkage

The first LSYPE cohort was designed from the outset to incorporate administrative data obtained from the National Pupil Database, held by DfE. It was one of the first studies designed in this way, and respondent's willingness to allow consent for their educational data to be obtained was a prerequisite for their inclusion in the study. DWP and DfE are currently arranging for data to be transferred to DWP for linking with DWP administrative datasets.

One of the key strengths of LSYPE, according to stakeholders, has been the inclusion of the National Pupil Database, which contains extensive information regarding students' educational experiences and allows for investigations of student academic outcomes. This dataset combines pupil-level achievement data with information held on pupil and school characteristics, sourced from Schools Census data (formerly known as the Pupil-level Annual Schools Census or PLASC). Many contextual factors were addressed by data contained in the Schools Census (in the case of schools) or collected during interviews (in the case of demographic or family background). This database was used as the basis for the original LSYPE study, and consent to link with these records was a requirement for inclusion as a participant. As such, due to the main focus of this study, and past importance in LSYPE, any future cohort of LSYPE should retain this inclusion criteria and should investigate accessing this data right from the beginning.

The first LSYPE cohort has demonstrated the benefits that linkage with administrative datasets can provide to survey data. However, should a second cohort take place further use of administrative datasets could aid in the development of a more broadly useful data source that takes into account additional areas that are known to impact upon young people's development. This desire for access to more data sources was commented upon during interviews with LSYPE data users, who highlighted that the first LSYPE cohort has been slow in linking with the Work and Pensions Longitudinal Study, for example.

Besides LSYPE, a number of other longitudinal studies have successfully managed to obtain permission to link into administrative records (see Table 8.1). What can be seen from these other studies are the range of potential data sources available, which could potentially be incorporated into a second LSYPE cohort. This information would allow

researchers the ability to more fully understand how different environmental and circumstantial factors influence an individuals future educational and economic potential and could greatly improve the utility of a second cohort.

Study	Cancer registry	Mortality (ONS)	Health	Education	Tax records (HMRC)	Benefit records (DWP)
ELSA			Hospital Episodes		Yes	Yes
MCS			Yes	Yes	Yes	Yes
NCD			Yes		Yes	Yes
ALSPAC	Yes	Yes	Yes			
US			Yes	Yes	Yes	Yes
LSYPE				Yes		Yes

## 8.2 Benefits of data linkage

There are a number of benefits to using administrative data which makes it attractive to researchers (Plewis et al, 2001). First of all, utilising data that has already been collected is cost effective in relation to obtaining such information directly from individuals, for example by reducing the need for costly interviewer time (Calderwood & Lessof, 2009). Secondly, a broader amount of information can be obtained on areas such as income or health, which respondents may not have an accurate recollection of, Information collected in this way may also be more reliable than information collected directly from respondents. A third benefit is the lack of intrusiveness and reduced respondent burden (Plewis et al, 2001), particularly for complex and uninteresting questions (Calderwood & Lessof, 2009).

In addition, making use of pre-collected administrative data has a number of practical benefits for collected survey data, principally for validating collected data and to correct for non-response (Calderwood & Lessof, 2009). Data obtained from routinely collected administrative datasets can be used to validate data collected directly from individuals and provide analysts with some indication of confidence for their conclusions. If shown to be reliable, the obtained data can then be used to minimise item non-response, either through direct substitution with the respondent-obtained data or by informing statistical imputation procedures. Another benefit, namely the ability to link to historical data collected before individuals were recruited into a study, is of particular relevance to LSYPE (e.g. by allowing researchers to model the effects of changes in schools or the level of school attendance prior to inclusion in the study).

Obtaining respondent consent for data linkage introduces additional considerations. Elements of bias may appear due to differences in the characteristics of people that provide consent, difficulties in obtaining matches when consent has been provided or because of bias in the nature of the data collected by administrative sources (e.g. excluding students from independent schools). However, given that all data collected can suffer from non-response including survey data, the amount of bias introduced as a result of data linkage maybe seen as trivial in relation to overall non-response (Calderwood & Lessof, 2009). The benefits of collecting extensive amounts of information that may not otherwise be obtained, or may be obtained only at a relatively higher cost, are arguably greater than any associated costs of obtaining consent.

### 8.3 Potential sources of data linkage for LSYPE

In this section, a number of potential sources of administrative data are described. Some of these have been used in the first LSYPE cohort (e.g. the National Pupil Database), or have been used in other studies (e.g. Health Episode Statistics). Others have been less widely used outside of their main fields of research (e.g. criminal records obtained from the Police National Computer). As part of this review, interviews with relevant departmental staff were conducted, which highlighted main issues and time frames in requesting access to such data.

Note, that as these databases are a historical record of individuals encounters with government agencies it is not necessary for consent to be obtained from wave one, as consent at later dates will allow retrospective inclusion of data. The one exception is the National Pupil Database, for which consent at wave one has been, and should continue to be a requirement for inclusion into this study.

One crucial point that is relevant for all potential administrative data sources is the need for DfE to consult with legal and ethical advisors to ensure that issues associated with confidentiality and informed consent are addressed. If a second cohort of LSYPE proceeds, this consultation will need to be completed prior to making requests for, or undertaking, data linkage.

### Geographic information

Three main sources of geographic information could be of conceivable benefit to a possible second cohort: the Index of Multiple Deprivation (IMD), the Child Wellbeing Index (CWI), and urban/rural classification. Such data could provide important contextual information regarding the communities that study members are growing up in, and the impact of local services provided by Local Authorities, NHS Trusts and other regional bodies. This information is easily accessible, and can be matched with an individual's postcode.

The IMD is made up of seven constituent domains (barriers to housing and services; education, training and skills; employment deprivation; crime; health and disability; income; living environment). Given the likely age of cohort members, the Income Deprivation Affecting Children Index is likely to be of particular use. This information is regularly updated and provided by ONS.

More recently, the CWI has been developed, in a similar format to the IMD and appropriate for small area estimation (Bradshaw et al, 2009). However, as a point of difference with the IMD this index has been created as a measure of wellbeing rather than of deprivation, although there are some similarities that can be drawn between the two. There are seven domains that make up the CWI: material wellbeing, health, education, crime, housing, environment, children in need.

Another widely used source of geographic information is the urban/rural classification. This was developed by the Department for Environment Food and Rural Affairs as a measure of the rurality of an area. There are two measures available: one for Census Output Areas, and a second for Local Authority region.

Such information is of importance in understanding the impact of the environment upon an individual's development. It has long been shown that people growing up in poorer communities on average under perform on a number of different criteria compared with those growing up in more affluent communities. Similarly, people growing up in more rural communities may face different social pressures from those living in urban areas.

## National Pupil Database

The National Pupil Database, maintained by the DfE contains information on each pupil within the education system of England, and the schools that they attend. Information contained within the NPD includes key stage assessment results and contextual data (such as schooling history, ethnicity, free school meals eligibility, and special needs status) obtained from the Schools Census data. Many of these measures are extremely useful for analysis relating to young people, and linkage to the NPD is therefore extremely important. However, these data do not cover privately funded schools, and therefore not all pupils will be covered by this dataset.

Applications require a business case to be lodged with the National Pupil Database and Dissemination Unit (NPDDU) in the Data Services Group of the DfE.

As a result of the relative ease of access and the data quality, this database has been extensively used by researchers. There is also a great deal of guidance provided by the PLUG User Group<sup>18</sup> hosted by the Centre for Market and Public Organisation at the

---

<sup>18</sup> <http://www.bristol.ac.uk/cmppo/plug/>

University of Bristol. Discussions with a member of the NPDDU suggested that data linkage for all possible members of the data frame maybe possible, but this would be dependent upon more specific information being provided.

Given that respondents are likely to be chosen from this database, the requirement that respondents allow access to their records for inclusion in the study, and the desire to maintain comparability with the first LSYPE, the NPD is likely to be the most important dataset for linking. A great deal of information is contained within this database that would not otherwise be possible to collect from respondents or would only be possible with great cost and additional respondent burden. This data would be seen as a minimum for many of the suggested analyses commented upon in an earlier section.

## Health Episode Statistics

The Health Episode Statistics (HES) are a collection of individualised health-related data collated by the NHS Information Centre (NHS IC) on topics such as health and lifestyle, hospital care, mental health and social care. This data is obtained from NHS trusts from around England and Wales. Of the three constituent datasets, only the Inpatient (including maternity) and Outpatient datasets are of a sufficient standard to be currently used for research purposes. Linkage of HES to LSYPE may still be extremely useful in analysing health issues affecting young people and how these relate to other background characteristics.

Typically, consent rates to access health data held on the HES have been quite high, with 79% of participants in the National Child Development Study between waves one and seven (Bhamra, Gatenby, Hacker, Killpack, Larkin, & Lessof, 2010), and 93% of participants at wave four of the Millennium Cohort Study (Gray, Gatenby, & Huang, 2010) being reported as providing consent. As such, it is reasonable to expect a sufficiently high level of consent amongst participants to warrant investigating access to this database.

Within the NHS IC there is a dedicated team setup to assist researchers with linking data held within these datasets. Additionally, mortality and cancer information can be obtained through the NHS instead of the ONS which had previously collected such information. Application can be made to the Medical Research Information Service (MRIS) to provide ongoing notifications in case of death or cancer diagnosis of named cohort members. Both ethical approval from the National Research Ethics Service and approval from the National information Governance Board for Health and Social Care's Ethics and Confidentiality Committee (NIGB ECC) are required. Approval from the NIGB ECC can reportedly take up to five months to secure. Data linkage is managed by the MRIS who require at minimum a full name and date of birth.

Health statistics would be particularly useful as a first step in documenting life course events that impact upon the individual. For example, the impact of Government supported initiatives on the health outcomes of 'at risk' youth could be investigated; the likelihood of their coming into contact with health services (e.g. drug and alcohol treatment, accident and emergency, mental health services) could then be compared with their peers considered to be at lesser risk. This type of information is likely to be more accurate if obtained from administrative records due to issues such as misreporting of complex health data or under-reporting of certain health problems. Given the benefits of having access to health data, and the relative ease and low cost of accessing these datasets, including the HES would be highly recommended.

## Taxation and benefit records

Possibly the database of most interest to researchers would be the Work and Pensions Longitudinal Study in which benefit and programme information held by DWP is linked with employment and tax information held by HMRC. This database therefore contains an extensive list of personal information regarding individuals.

Taxation and benefit records have been widely used in longitudinal and other large scale studies, and provide useful economic context which can be difficult to obtain from surveys alone. Typically, such information has been requested from parents and caregivers of young people taking part in studies, as it is parental income that exerts an impact on their child's development and access to resources. However, such information is becoming increasingly difficult to obtain due to increased security procedures.

While data relating to income and benefits can be directly asked of respondents, it can be a sensitive topic that people maybe unwilling to discuss in the setting of the interview or maybe unwilling to detail directly to the interviewer. Furthermore, for some people there maybe a degree of misreporting for reasons such as the complexity of their financial affairs (e.g. income may consist of benefits and child support or they may not be aware of the exact benefits they are claiming) or because the individual spoken to is not responsible for household finances. Another advantage that can be seen of using taxation and benefit records is the ability to identify when people stop and start employment, a task that can be difficult when asked of respondents directly due to recall errors.

In discussion with DWP it was mentioned that a major issue that would need to be negotiated should data linkage occur if a second cohort take place is the identification of a sponsor within DWP who would be responsible for managing the application process and eventually data provision. This person needs to be identified when a formal request is made for access to data sets to ensure the interests of DWP are met.

Attempting to obtain access to such information is highly recommended for a second cohort of LSYPE. This information will maximise the benefit of the information obtained by LSYPE for little cost, and will increase its utility for other Government Departments that maybe interested in young people. Furthermore, it maintains comparability with the first LSYPE cohort as well as adding to the income data that is already intended to be collected directly from respondents.

## Police National Computer

The Police National Computer (PNC) is maintained by the National Policing Improvement Agency (NPIA) with regular data extracts being provided to the Offender Management and Sentencing Analysis Services (OMSAS) of the Ministry of Justice (MoJ) for research use and to provide to academic users. Information gained from contact with the MoJ recommended use of this extract, which contains identifiable information including a unique PNC number for long term tracking. These data would be a valuable addition to LSYPE because they can provide information on young people's offending to add to that obtained from LSYPE in terms of risky behaviours and criminal convictions.

Information regarding young people's involvement with the criminal justice system can be seen to be relevant to issues of crucial policy importance, and would increase the utility of the LSYPE dataset for policymakers outside of the DfE. For example, such information could be utilised to investigate background factors relating to young peoples later criminal behaviour, or used to evaluate changes in behaviour for young people considered 'at risk' over time. Such information, if asked directly from respondents is likely to result in a great deal of under reporting of such activities and as such data collected from the PNC can be seen as having greater reliability.

Access would require approval from the NPIA's PNC Access Panel (PIAP) who administer the database, and the MoJ who will provide the data and arrange linkage. If approval is granted, a data sharing agreement and privacy impact assessment would need to be completed and submitted. Details regarding how the data are stored, who has access to such information and the length of time such information is held would need to be negotiated. These considerations are seen to be of particular importance to the MoJ.

Discussions with members of the NPIA and MoJ regarding the possibility of data linkage occurring for all members of the sample frame led to several concerns being raised. In particular the issues expressed focussing on how such data linkage could be completed while respecting legal and ethical requirements of access to an individual's data without signed consent.

## Higher Education data

The Higher Education Statistics Agency (HESA) is the agency that undertakes the collection and analysis of quantitative data on behalf of the Higher Education sector. Of the data records held by HESA, two would be of potential use to a future LSYPE cohort. The Student Record dataset contains information related to students' current higher education (e.g. course type, funding) and is collected annually, and the Destination of Leavers dataset contains information related to employment and earnings upon completion of higher education. Generally both datasets are of high quality, but the Destination of Leavers dataset typically receives 70-80% response.

Given plans to increase the education and training participation age for young people, and ongoing interest in young people's progression into higher education and post-compulsory education and training, the use of datasets such as the Student Record and Destination of Leavers is likely to be of much interest to policy makers. For example, this would be used to understand and evaluate policies developed to increase the participation rates of young people from disadvantaged backgrounds in higher education. Currently, such information can only be collected retrospectively and the type of information that can be collected in such a way is vastly more limited.

Within HESA, the Information Provision Team manages and action requests for data linkage. Data linkage is managed from within HESA. Applications can be made via HESA website ([www.hesa.ac.uk](http://www.hesa.ac.uk)) and require notification of what the data will be used for. When this application is made HESA make contact regarding cost and timings of data provision.

Linking to all possible individuals in the sampling frame is dependent upon anonymised records being provided to the research team; the provision of anonymised records is the norm for data provided by the HESA. However, it is likely that identifiable information could be obtained if requested by the DfE.

### 8.4 Respondent views on data linkage

Respondents in the first LSYPE cohort were asked about their views on data linkage (in relation to confidentiality) during interviews. Respondents were confident that the information they gave was treated in 'strict confidence'. To them this meant that their answers were not:

- Stored with their names and addresses;
- Passed to 'third parties' (such as market research agencies); or,
- Used for anything other than research purposes (such as marketing).

Respondents had no concerns about their survey answers being linked to their exam results as currently happens. The latter were seen as being a matter of public record and

for some were something to be proud of, so other people having access to them was not of concern.

Respondents were asked about how they would feel about requests to link their survey answers to health or tax and benefit records. Views were mixed, reflecting perceptions of how sensitive these data are. In terms of health records, there was a feeling that it might put some people off taking part, though no respondents stated they would be unhappy giving their consent to this request. One respondent stated that he would not mind providing this information but he did not have any health conditions, and that if he did, he might feel differently.

Having access to tax and benefits records was considered to be problematic because the information was personal or sensitive and there could be more serious consequences if the information got into the 'wrong hands'. One respondent thought that when young people are first asked to take part, when they are at school, they may not mind as they would not be paying tax or receiving benefits. However as they got older they might become more suspicious about it and it could prompt them to drop out.

These concerns would have to be considered (and allayed) among respondents and their parents, should data linkage be pursued.

## 8.5 Recommendations

Overall, data linkage with data already held by Government Departments has been shown to be an effective strategy to increase the utility of data collected directly from the respondent. Principally, it has been found to save both time and money, as information that is already collected does not then need to be obtained from the respondent directly. While there are barriers to accessing such information (e.g. the need for signed consent from respondents, security arrangements required by the data holders), such barriers have been successfully overcome in the past.

Relative to collecting data during interviews, obtaining administrative data is less costly, and allows other important information to be obtained during interviews that cannot be obtained elsewhere. This increases the utility of data sources to researchers and policy makers. Other benefits include allowing for the validation of data collected through interviews, reducing respondent burden by accessing information already held about the individual, and informing non-response patterns.

NatCen recommends that data linkage in a second cohort should be considered and that the following should be taken on board:

- Data linkage with the National Pupil Database occurs within the first LSYPE cohort, and NatCen recommends that this should continue if a second cohort took place.

- Data linkage with administrative data sources is a cost effective and valuable method of adding to the utility of a possible second cohort. Should this occur, promising sources of data are the National Pupil Database (NPD), Health Episode Statistics (NHS) and the Work and Pensions Longitudinal Study (DWP). These datasets have been previously accessed by other longitudinal studies, although these have not had the same focus as LSYPE (e.g. MCS). The data obtained from linking to these datasets is likely to increase the utility of LSYPE to researchers and policymakers in a number of different fields.
- Other sources of administrative data should be considered and discussed. These include criminal records held on the Police National Computer and higher education data. These datasets are likely to be the most difficult to access due to the sensitivity of the data held or are likely to be of the least immediate use/benefit. Therefore NatCen would recommend that further consideration be given to using these sources only after further discussion with the relevant data owners and a clear rationale for their inclusion is made.
- When exploring possibilities for data linkage, several Departmental representatives mentioned the legal requirements that would need to be satisfied and the need for a Departmental sponsor both within the DfE and in the Department controlling the administrative dataset. If a future LSYPE took place, these sponsors would need to be identified, as would any legal basis that would allow data sharing to occur.

## 9 Funding

Although some of the benefits of funding a second cohort are outlined in this report (notably in the second Chapter), a final decision on funding a second cohort study cannot be made until after the full implications of the autumn Comprehensive Spending Review have been considered.

This Chapter explains the history of funding LSYPE, originally intended to be a cross-departmental funded study, and looks at the lessons learned. It reviews some other cross-Government/organisation funded studies as well as outlining some of the potential disadvantages when following such a strategy. Ultimately DfE will need to take the final decision on the strategy it wishes to pursue, however this Chapter makes some recommendations on the best way to obtain cross-departmental funding should this be the direction followed.

### 9.1 Funding of LSYPE

Funding was allocated across financial years rather than waves of the study because one wave covers more than one financial year.

The original study was initially and mainly funded by HM Treasury. The Longitudinal Surveys Team at DfE applied for funding for the study in the 2002 and 2004 Spending Reviews which covered 2003-08. In addition, DfE has used funds from their central research budget. DfE has funded all Waves, and these costs cover the data collection, fieldwork and additional data enhancement costs.

### 9.2 Cross-departmental funding and lessons learned

LSYPE was designed to be a Government wide study which received initial funding direct from HM Treasury. As time went on DfE ended up managing and funding the bulk of the study and as the cohort aged, they moved out of DfE target policies into those of BIS and DWP, who became co-funders.

The Machinery of Government changes in 2007 negatively impacted on obtaining funding for LSYPE from other Government Departments, as the remit of the Department became much wider in the change from the Department for Education and Skills (DfES) to the Department for Children, Schools and Families (DCSF). Given the expanded remit of the DCSF, there was little incentive for other Government Departments to fund the study. Consequently, DCSF experienced difficulties in getting other Departments such as the Department of Health to sign up to LSYPE, since the health of young people was now within the remit of both Departments.

From stakeholder interviews it was reported that the first LSYPE cohort was too narrowly defined at the outset, so that other Government Departments were less interested in co-funding from the start. It was not clear how the study was relevant to other Departments or what opportunities it presented. Combined with a lack of awareness of the study and the wider remit of the new DCSF, this also meant that other Departments were not keen to fund later waves of the study as they were unsure of the benefit for their investment.

### 9.3 Other cross-Government funded studies

There are a number of other cross-Governmental funded studies. Although many of these studies have different funding strategies, as they were designed with different audiences in mind, this section looks at these studies in order to look at the range of possible funding options and to highlight options that may be useful for a second LSYPE cohort.

Understanding Society is mainly funded by the Economic and Social Research Council (ESRC), an Arms Length Body of the Department for Business, Innovation and Skills (BIS). However, other Government Departments have also provided funding. For example, the Department for Communities and Local Government (CLG, who are interested in housing, labour market and occupation, neighbourhoods and social networks) provided funding in return for access to the datasets being provided with analytical reports relevant to the organisation, and gaining input into the design of future Waves of the study.

The British Social Attitudes survey is mostly funded by the Gatsby Charitable Foundation (a Sainsbury family charitable trust), but also receives funding from the ESRC, the Nuffield Foundation and Government Departments, including DH, DWP, DfE, BIS and the Department for Transport (DfT). Contributing Departments have funded their own question modules within the survey, such as the DWP designing and cognitively testing questions relating to people's attitudes to disability and child poverty in recent years.

The Millennium Cohort Study is largely funded by the ESRC, but has also obtained funding from the Wellcome Trust and from a consortium of Government Departments led by the ONS, including the DWP, the DH, DfE and all the devolved administrations (the Welsh Assembly Government, Scottish Government and Northern Ireland Executive). The funding consortium decides the content of the study, and there is a strong focus on socio-economic data which has meant that DWP have been quite heavily involved.

The English Longitudinal Study of Ageing is funded half by NIA (the US National Institute on Ageing) and the other half by a consortium of Government Departments coordinated by the ONS, including the DH, the DWP, HMRC, DEFRA, CLG and the DfT.

## 9.4 Disadvantages of cross-departmental funding

Ensuring cross-departmental funding is clearly beneficial in that it reduces the burden of cost for the Department. This may be especially prudent given the current financial climate. However it is important to note that this approach also has some potential disadvantages.

One of most likely consequences of pursuing a strategy of cross-departmental funding will be the requirement to broaden the scope of the study to make it more attractive to other departments. One approach would be to increase the length of the survey so that additional questions could be included. However this could lead to substantial increases in the overall costs and place additional burden on the respondent increasing the risk of attrition. Alternatively, other departments could include questions at the cost of some of the questions intended for inclusion by DfE. This also has a number of potential disadvantages. It could lead to a more diluted study in which there is less focus on young people's transitions through the education system. This in turn may have consequences for the compatibility with the first cohort detracting from potential comparative studies for evaluating change. Finally, there is the additional risk that LSYPE becomes an omnibus type survey in which different questions are asked in each wave of the study. This could have severe implications for its strength as a longitudinal study whereby questions are repeated over time to enable an understanding of development and change.

This is not to preclude the idea of pursuing cross-departmental funding, but to give some forewarning of the potential issues that may arise that will need to be adequately managed. Decisions regarding which questions will need to be retained will need to be well thought through, particularly in ensuring comparability with the first study. Other departments should be made aware of what constitutes a longitudinal study and the importance of repeating questions over time.

## 9.5 Recommendations

It is clear that other longitudinal studies have secured cross-departmental funding, and despite the notion that Government funding may be difficult to obtain in the current fiscal climate, there are lessons that can be taken forward from this approach that will help look at ways of doing this. Should DfE decide to pursue cross-departmental funding NatCen would recommend the following:

- In the current climate of financial prudence, the distinctiveness and unique benefits of LSYPE need to be clearly defined and communicated to other Departments at the outset, even if these benefits may not be immediate for all potential funding Departments

- Cross-sectional analysis of the first wave of the study should be encouraged both internally in the Department and externally, which would help to engage potential funding Departments early on in the life of the study.
- Most of the cross-departmental funding in the examples outlined above was secured by ensuring that the contributing Departments have a say in the questions asked and in some cases receive dedicated reports summarising relevant analyses. Such an approach could be taken to target Government Departments and would involve broadening the scope of the study, for example to include more work-related or health-related questions.
- A word of warning however. There are risks associated with broadening the scope of the study which could lead to a diluted study with less focus on young people's transitions and reduce its comparability with the first LSYPE as well as its potential as a longitudinal study. Points which were also raised at the LSYPE User Group Meeting. If the Department decides to pursue cross-departmental funding it is vitally important that this is carefully managed and the integrity of the study design maintained.

# 10 Communication plan

The structure of this Chapter differs to other Chapters in this report because this is the first time a communications plan has been developed for LSYPE. Evidence to support the findings is drawn from workshops and interviews conducted with young people who represent future respondents; young people who participated in the first LSYPE cohort, stakeholders and staff. More details of the suggested communication strategy can be found in Appendix M.

Historically, no communications plan has existed to govern how LSYPE is communicated to young people and their parents to boost participation and retention. Furthermore, budget has not been ring-fenced to fund adequate resource to implement the communications plan and develop a set of professional and engaging communication materials. Within the context of today's falling participation rates, and the reality that competition for young people's attention by brands and initiatives is greater than it has ever been, it is important to explore the role that expert communications can play in getting the LSYPE onto young people's radar and making it feel important and interesting enough to encourage them to take part.

The first LSYPE cohort data and findings have not been systematically communicated to all potential users (the data is made available through the iLSYPE website), so awareness of the LSYPE, its potential use and overall value is patchy among important audiences such as policymakers, and virtually non-existent among organisations who could also benefit (such as community groups). Given the current economic climate, it is important that a second cohorts maximizes its return on investment– meaning that as many people as possible should be able to access, interact with and benefit from the data and insights. Furthermore, the value for money for the first LSYPE cohort could be maximised by encouraging people to think about comparisons between the two cohorts – an objective that can be met through the development of a communications plan.

## 10.1 Communicating with respondents

The target audience in terms of potential respondents and people that will influence their participation in a second cohort are:

- Young people aged 13, to encourage them to participate in the study;
- Young people aged 14-18 who are already participating, to encourage them to stay involved with the study;
- Parents of young people who are invited to participate in the study;

For this review, four workshops took place with young people to ascertain their views on taking part in a longitudinal survey such as LSYPE. The primary insight was that young people would take part in such studies but with a degree of apathy: 'why not?' They didn't appear to be motivated that much either way. Therefore, if they were asked in the right way and taking part is made simple, young people are likely to take part. Convenience was found to be the most dominant factor over incentives, impact of findings, making a difference and peer endorsement.

Secondary insights from workshops included:

- *Young people value transparency:* Many young people displayed keen interest in the study's purpose, who will use it and who has commissioned it.
- *If they're going to bother doing it, make it matter:* The general mood was that young people wouldn't want to go to the effort of doing it if it wasn't going to make a difference to people's lives, so regularly keeping them up to date with what insights are being drawn and how these are being used is important.
- *Inclination to participate declines with age and possibly varies according to gender:* 13 year-old boys are much more likely to agree to participate than 15 year-old boys; boys seemed to need stronger persuasion to take part than girls.
- *They are aware of issues about data security but concerns are easily allayed:* They are aware that personal information needs to be safeguarded but were easily satisfied when told the study is safe.
- *They're not interested in the results for results sake:* However, they are interested in knowing how they compare with their peers.
- *Peer endorsement may be important for some:* Girls appeared to be more anxious to know that other young people were taking part in the survey.

## Interviews with former respondents

Findings from interviews with LSYPE respondents suggested that young people have a range of motivations for taking part and prefer to be interviewed using a number of modes and locations (at home, at school, face to face, online etc). It was also found that parents can act as gatekeepers; particularly with younger age groups and that they may have more influence over the young person's decision to participate in the study initially than in later Waves.

The research highlighted that the following issues are important in terms of how longitudinal studies should be communicated to young people.

Tone of voice/information conveyed: This should be informative, simple and open and be clear to respondents this will have minimum impact on their lives. It should outline who is responsible for the study and what it is going to be used for.

Branding: The study should have a brand and this should stand for transparency, convenience, simplicity and making a difference.

Messaging: A range of messages regarding a study will resonate with young people, depending on their personality and motivations for taking part. These include:

- This won't inconvenience you;
- This is the largest survey of its kind in this country;
- This is your chance to tell politicians and people that influence your lives what you think about important issues;
- Your opinion will make a difference to young people's lives;
- Thousands of young people like you take part every year – by taking part you can find out what they think about things too;
- You will be rewarded financially; and,
- It is safe to take part.

These messages are not ranked by degree of influence or importance on young people. The exact wording, the degree of importance to young people and mode of communication used to get these messages across to potential respondents should be developed and tested with young people before being rolled out. Messages for parents should also be developed and tested<sup>19</sup>.

Incentives: The use of incentives has already been discussed in Chapter Six. The research found that money is the best option as an incentive because the costs attached

---

<sup>19</sup> Please note that communications for parent and teachers was outside the remit of this review

to providing a range of incentives from which young people can choose may be prohibitive.

## Communicating with Respondents - Recommendations

The research with young people has led us to a number of strategic recommendations to incorporate into a communications plan to boost participation and retention:

Carry out research among parents: there is a need to explore parents' roles as gatekeepers and possible motivations and levers to encourage their participation. This insight should be fed into the development of communication materials for parents and interview methods.

Inform widely and regularly: A second cohort needs to be promoted directly to young people, and also to those who influence young people e.g. parents. Materials need to be developed for all audiences, with regular updates on the cohort's progress and insights that have been gathered to date – appealing to people who need to feel their efforts are worthwhile, have altruistic moments or who want to learn something. A range of materials should be developed to facilitate this, including a professional, intuitive, simple website, text messages and a newsletter that can be distributed by post or email.

## Channels of communication

The following are suggested as key channels that can be used to communicate with respondents about a second LSYPE cohort:

Website: The study website could become the study 'hub' over time with respondents returning to it to claim their incentives, view the latest findings and see where the study has appeared in the media.

Information packs for young people and parents: The packs could contain a range of materials designed to answer questions about the study and inspire people to take part including:

- Booklet - Evidence from other surveys suggests that an attractive, engaging and inspirational booklet reinforcing the purpose, importance and uniqueness of the study can create a positive perception of the study in the minds of the respondent and encourage them to take part. Interviewers can also have extra copies to hand out on the doorstep.
- Membership card - To allow young people to register online, update their details and claim their incentives. This card could contain details of the study, website and contact details and a respondent ID number.
- Parental letter of consent.

Emails and text updates: Short, plain text based email updates could be sent to young people regularly to keep them up to date with the study, and drive traffic back to the website after the initial interview. With an updated friendly and interactive website to go to young people may find this a less intrusive way to stay in touch between Waves. They should also been given the opportunity to opt out of these updates.

## 10.2 Distributing findings and engaging users

A second aim of the development of a communication plan was to identify how to increase the number of people who engage with LSYPE findings, especially if a second cohort was to be undertaken. Research users can be both direct and indirect:

- Direct (meaning people who actively seek out data and are used to interrogating data/findings) e.g. academics, analysts, policymakers, media.
- Indirect (meaning people who do not actively seek out data/findings but who would benefit from the insights in their work) e.g. youth and community groups, police, local councils, schools, primary care trusts.

Key findings from the User Group Seminar included:

### Awareness and knowledge of LSYPE and the value of longitudinal studies is patchy

- Awareness levels of LSYPE were variable. Policy makers generally did not realise that LSYPE collected information from parents.
- Policy makers and analysts tended to have limited knowledge of what others were doing with the data.

### There is a need to provide data and report on findings in a range of ways

- Policy colleagues leave it to analysts to interrogate data and prefer face to face briefings/presentations.
- Nearly all User Group Seminar respondents proactively looked for research themselves and favour a mix of face to face (informal chats and briefings), emails, publications and websites.
- Some User Group Seminar respondents felt that data outputs and datasets should be marketed to make them easier engage with.

### Findings should be shareable

- Nearly all data users share research with colleagues via word of mouth or emails to colleagues with links attached, so there is a need to present findings in a format that people can readily share with each other (e.g. in an email format with embedded links leading back to further info on the website, or in PowerPoint presentations that people can give to their colleagues).

### Data should be easy to use

- Some felt that the LSYPE data sources are complex and whilst the data are well documented, they are very difficult to use, simply because of the size of the datasets and the number of variables they contain.
- A subgroup of LSYPE data users unanimously said that the data would be easier to use in future with more derived variables, particularly for the life history file which is very difficult to use.

### Consider creating clarity around the study brand

- Some User Group Seminar respondents were confused about the two names for the study. LSYPE is generally used to describe the study within the Department and amongst 'expert' research users though it is known as Next Steps by respondents. Some did not know that the two names relate to the same study. While not critical to success, a single name for the study would make it easier and more cost effective to build the profile of the study among all target audiences, particularly through the media and online. Respondents in particular would benefit from recognising media coverage (in print or online) about the study, as it would help them understand its importance and how it's making a difference - they won't recognise any media coverage about LSYPE if they only know it as Next Steps. If the Department were to adopt a single name/visual identity for the study, it will need to be appropriate for all audiences.

## Communicating with users

To promote access to the findings from studies such as the LSYPE a main priority is to build the preparation, release and distribution of study findings for public consumption into an overall study communication plan. This requires a budget for the production of materials and the resource required to communicate study findings from the outset. The budget will depend on the precise nature of materials to be developed and whether production will be resourced in-house or through an external supplier. The first step is to produce a specification or brief for materials upon which cost estimates can be based.

Findings should be shared in a range of formats, and tailored as much as possible to different audiences. When possible, findings should be supplied to intermediaries such as the media, other Government Departments, Strategic Health Authorities, trade associations, sector representatives (such as the National Council for Voluntary Organisations) and so on, to encourage them to cascade the findings throughout existing networks and thus disseminate findings more widely than only through DfE channels.

NatCen and Public Zone recommend appointing a communications manager to manage the packaging and distribution of study findings, as well as communication to respondents and teachers, as described earlier in this Chapter. This ensures that

communication is coordinated, consistent and adequately resourced, while minimising costs by keeping central resources in-house. A suggested delivery plan, incorporating both communication to respondents, those that influence respondents and users, is outlined in the final section below.

### 10.3 Communication Plan - Recommendations

#### Communicating with users and respondents

- Appoint a Communications Manager – recruit a full time, expert in-house communications manager. Their role would be to:
  - offer strategic advice about communications to young people and parents
  - develop, share and coordinate the delivery of the communications plan
  - liaise with partner organisations to agree roles
  - develop materials
  - ensure key messages are communicated consistently and
  - evaluate activity against success criteria.

A benefit of this role is that their expertise and knowledge will stay within the Department and costs will potentially be lower than outsourcing the role to an external agency.

- It is best practice to review communication regularly - at the very least there should be an annual review, but traffic to study websites should be monitored monthly and response to emails and SMS messages every time they are sent out. Insight can then be used to hone materials and optimise the impact of communication. It is also advisable to track who is accessing the findings, in what format and how often in order to gauge which formats are most popular with research users, which audiences are most readily engaging with the study and how communication can be improved to maximise interaction. A plan for effective monitoring and evaluation should be developed from the outset.

#### Communicating with respondents and those that affect them

- Prepare a communications plan. This is the communications 'bible' for the communication managers featuring target audiences, branding and key messages, materials to be produced with costs, timeline for distribution and measures for success. The benefits of this is that it ensures coordinated and consistent approach across multiple partners.
- Develop and test materials for parents and young people. Once a communication plan has been developed and the range of required supporting materials identified, this can move onto the production phase. Huge value is derived from

testing messages and communication materials with young people before rolling them out so this is strongly recommended. This also promotes confidence that communication materials will deliver greater return on investment.

## Communicating to study users

- Prepare a database of report/data recipients – what areas they work in, what policy areas they are interested in and in what format they prefer to receive findings. This database should cover a broad range of audiences. Typical users of the LSYPE findings and ‘non analyst’ audiences (i.e. people who are not expert research users but whom could derive huge value from the study findings in their work including youth workers, teachers etc). This database should be updated and cleaned regularly throughout a study’s duration. This enables findings to be packaged and distributed to recipients in their preferred format, which should boost usage.
- Website – develop a single, professional looking and usable website (currently there are several for the LSYPE which should be consolidated). The website could be used to present the findings in a range of ways (e.g. downloadable reports, presentations, data visualisation and links to the UK Data Archive where the raw data is held) to encourage visitors to use and share the findings and give people the opportunity to register for email alerts when new findings are posted to the website. The website could be developed on the same platform as the iLSYPE website with a caveat that technical or confusing language should be avoided and the site needs to be simple to use and search. This creates a single, user-friendly platform that people can be directed to, to interact with findings.
- Prepare a distribution plan and timeline - develop a timetable for proactive communication of findings, developing templates for a range of formats:
  - Email bulletin (people can sign up to this via the website and share it with their peers).
  - PowerPoint presentations for policy briefings.
  - Downloadable printable and shareable reports.
  - Media releases/briefings (targeting specialist and sector publications favoured by the target audience).

The benefits of this are coordinated consistent communication that can be staggered throughout the year, a centralised plan and templates that can be followed should resourcing change.

# 11 Conclusion

This report reviews the first LSYPE cohort, and has recommended a number of improvements to the design that would enhance the setting up of another longitudinal study of young people, whether a second LSYPE cohort or another similar study.

Findings from a wide range of research methods, including depth interviews, user-groups, telephone interviews and desk research have been combined in a unique way to address the key objectives specified by DfE. Recommendations have been generated by drawing upon the expertise of key methodologists and stakeholders (across Government and academia).

## 11.1 Objectives of the review and how these have been fulfilled

**Objective 1 – Critically appraise the methodology for the first LSYPE cohort to inform the design of a second LSYPE cohort**

NatCen recommends that the new cohort should begin when pupils are in Year 9 to maintain comparability with the first cohort. The Annual Schools Census should be used as the sample frame for a second cohort. A clustered sample is then recommended, preferably by school, although other possibilities include pupils' postcode sector or super-output area. The overall sample size for a second cohort should be at least as large as that obtained in the first LSYPE cohort. Better still it should be large enough to achieve an adequate effective sample size for each important subgroup in later waves. Specific subgroups of interest should also be boosted. In terms of mode, NatCen recommends that interviewing should be face to face at Waves 1-3 and then a mixed mode sequential design introduced at Wave 4. After analysing response rates it is clear that reissuing non-responders in subsequent waves could improve response rates in a second cohort and that hard to reach groups should be specifically targeted if the resources are available (see below). In order to properly monitor changes underpinned by the Education and Skills Act 2008 (which increased the age of compulsory education to 18) a second cohort should begin in 2012 or later. This would capture the experiences of a cohort of young people that fully experience the impact of the changes brought about by the Act. NatCen recommends that the timing of interviews remains the same as in the first LSYPE cohort (spring/summer), but that care is taken to avoid exam periods. Research into incentives used in the first LSYPE cohort and other studies suggests that unconditional incentives should be offered from the first Wave of data collection.

### **Objective 2 - Review of how LSYPE has been used both within DfE and externally**

The LSYPE has been used extensively by Government Departments and academics, both within the UK and internationally. With a dedicated communications strategy, a second cohort would be able to build on the reputation of the first, and would be used even more widely. This would also provide an additional return on investment, because the two cohorts would complement one another – comparative studies using the two datasets would ensure that the first cohort also remained a valuable resource, as well as maximising the impact of the second cohort. Increasing awareness of the LSYPE, including dissemination of current findings, would encourage further use – recommendations for a communications plan are outlined under objective 7.

### **Objective 3 – Identify evidence gaps that a second longitudinal study could uniquely fill**

The review clearly highlights the uniqueness of the first LSYPE cohort. There is a wealth of information collected in the LSYPE that is not available in other studies, particularly concerning young people's transitions to adulthood, their attitudes and aspirations and their relationships with their parents and peers. The review also identified additional evidence gaps that a second cohort could fill by including additional questions on health, psychological and personality measurements, school choice, career aspiration, parenting, social networks (friends, siblings, contact with absent parents etc) and sexual orientation.

### **Objective 4 – Review how LSYPE has been funded**

Should a second LSYPE cohort occur it would be imperative to strike a careful balance between widening the scope to interest other Departments without broadening it to the extent that it loses its original focus. Should DfE opt for cross-departmental funding NatCen recommends that contributing Departments and organisations should be given opportunity to feed into the design of the questionnaire and in some cases receive dedicated reports summarising relevant analyses. This could maximise interest in the LSYPE and assist to embed funding sources, as the data becomes an invaluable resource to a number of Departments. Cross-sectional analysis of the first wave of the study should also be encouraged both internally in the Department and externally, which would help to engage potential funding Departments early on in the life of the study.

### **Objective 5 – Explore potential for data linkages with other datasets**

NatCen recommends that data linkage be pursued with the National Pupil Database, Health Episode Statistics and the Work and Pensions Longitudinal Study. Other sources of administrative data could also be considered, however these three should be the priority source of data linkage.

### **Objective 6 – Make recommendations on an analysis plan for the second cohort**

A second cohort would be invaluable, providing data with which to explore changes in the experiences and context of young people's lives within a newly emerging set of policies, particularly through comparisons with the first LSYPE cohort. The broad range of contextual information collected by the LSYPE could be used to develop a better understanding of differences between young people, as well as the factors that contribute to these differences. NatCen recommends the exploration of Key thematic policy strands, which include: young people's attainment; school choice; school curriculum; improving the achievement of disadvantaged young people; parental engagement; transitions into further education, training or employment; vulnerable young people; vocational qualifications; young people's engagement in risky activities; and the role of the Big Society (teachers, parents, communities) in helping young people develop.

### **Objective 7 – Make recommendations for a communications strategy**

A communication strategy for the second LSYPE should target two main groups. Firstly, it should encourage young people to participate in LSYPE initially and ongoing, and secondly, it should increase the number of people who engage with the study findings. To improve communication with users and respondents, NatCen recommends that a communications manager be appointed. A communications plan should be prepared and materials for parents and young people developed and tested. NatCen recommends that this communications plan includes a database of report/data recipients, the development of a single, professional looking and usable website and a distribution plan and timeline.

## **11.2 Essential and desirable elements of a new LSYPE cohort**

This review was commissioned in a very different financial environment from that currently in place as it goes to publication. It is therefore possible that if a second LSYPE cohort is commissioned, decisions may have to be made concerning the elements that are the most important to include and which, although they may be desirable, will no longer be feasible in the current financial climate. Table 11.1 illustrates the essential recommendations NatCen has made for a second LSYPE cohort, alongside recommendations for what might be excluded if this was to go ahead with a reduced budget. It is difficult to know what scale of budget reduction might occur, so the recommendations have been devised on the basis of a relatively minor reduction. Any major budget reduction would require revisions to the study design, and has therefore not been examined here. Any reduction in funding for a major longitudinal study would have to ensure that a careful balance is struck between conserving resources and losing vital depth to the study.

**Table 11.1 Recommendations for a second LSYPE cohort**

Topic	ESSENTIAL recommendations	DESIRABLE recommendations
<b>Rationale for a second LSYPE Cohort</b>	<p><b>Analysis plan</b></p> <ul style="list-style-type: none"> <li>• Focus on the particular strengths of the dataset, drawing out what makes the LSYPE unique from other sources of data.</li> <li>• One of the most obvious strengths would be the possibilities that a second cohort offers for comparative studies with the first LSYPE cohort. Use this to explore change in the experience and context of young people's lives and enable an evaluation of Government policies.</li> <li>• Use longitudinal analysis to explore young people's development over time, to gain a better understanding of the different pathways they follow.</li> <li>• Explore thematic policy strands such as: young people's wellbeing; young people's attainment; young people and school choice; improving the achievement of disadvantaged young people; parental engagement; transitions into further education, training or employment; vocational training; vulnerable young people, young people's engagement in risky activities and the effect of the Big Society (school, parents, community) on supporting young people to make successful transitions.</li> </ul>	<p><b>Content</b></p> <ul style="list-style-type: none"> <li>• Improve the content further by including information with a stronger focus on health, psychological and personality measurements, school choice, career aspirations, parenting, social network information (friends, siblings, contact with absent parents etc) and sexual orientation.</li> </ul>
	<p><b>Advisory Panels</b></p> <ul style="list-style-type: none"> <li>• Two Advisory Panels are set up for a new LSYPE cohort. The first being an Expert Advisory Panel consisting of stakeholders including academics and policy makers, and the second being a Young Person Advisory Panel.</li> <li>• Advisory Panel members should be consulted on issues such as questionnaire content, use of incentives, methods of keeping in touch with respondents, the design of the study and the best ways of presenting study literature.</li> </ul>	

**Table 11.1 Recommendations for a second LSYPE cohort continued...**

Topic	ESSENTIAL recommendations	DESIRABLE recommendations
	<p><b>Maintained schools</b></p> <ul style="list-style-type: none"> <li>• A two stage sample of maintained pupils through the Annual Schools Census should take place.</li> <li>• Schools should be used as the Primary Sampling Unit (PSU).</li> <li>• Following the first LSYPE cohort, approximately 30 pupils per PSU should be chosen.</li> <li>• PSUs should be stratified before sampling. Recommended stratifiers include: type of school, GOR, urban/rural or academic performance.</li> <li>• Pupils could be stratified by variables such as ethnicity, gender, FSM, IDACI or KS3 achievement.</li> </ul> <p><b>Independent schools</b></p> <ul style="list-style-type: none"> <li>• A two stage sample of independent pupils using the school as the primary sampling unit (PSU) is recommended.</li> <li>• Following the first LSYPE cohort, approximately 30 pupils per PSU should be chosen.</li> <li>• PSUs should be stratified before sampling. Recommended stratifiers include: whether they are single or mixed sex, GOR and academic performance.</li> <li>• Independent school pupils do not need to be further stratified.</li> </ul> <p><b>Recommended age of cohort and fieldwork year</b></p> <ul style="list-style-type: none"> <li>• The age of cohort in a second LSYPE should be Year 9 pupils.</li> <li>• Fieldwork should not commence before 2012.</li> <li>• The timing of interviews should remain the same as the first LSYPE cohort with interviews taking place over the spring and summer months.</li> </ul>	<p><b>Maintained schools</b></p> <ul style="list-style-type: none"> <li>• Pupils from certain subgroups of interest should have an increased selection probability.</li> <li>• Sample sizes should not be chosen to achieve a large Wave 1 sample, but be chosen to achieve an adequate effective sample size for each important subgroup in later waves.</li> </ul>

Table 11.2 Recommendations for a second LSYPE cohort

Topic	ESSENTIAL recommendations	DESIRABLE recommendations
Response rates		<ul style="list-style-type: none"> <li>• Reissuing non-responders in subsequent waves could improve response rates in a second LSYPE.</li> <li>• Hard to reach groups should be specifically targeted, such as those with lower educated parents and lower socio-economic status, those from lone parent households and those living in disadvantaged circumstances.</li> </ul>
Mixed Modes	<ul style="list-style-type: none"> <li>• The first three Waves should be conducted using face to face interviews with young people.</li> <li>• A sequential mixed mode approach is recommended after Wave 3 with the cheapest mode used first and then more expensive modes worked through for non-responders.</li> <li>• The questionnaire survey should be designed from the outset to be mixed mode, and be designed so that questions with the greatest risk of measurement error between modes are not asked in mixed mode Waves if possible.</li> <li>• Internet interviewing should not be a sole option, as some respondents will not have access or literacy.</li> </ul>	
Use of Incentives	<ul style="list-style-type: none"> <li>• Unconditional monetary incentives of approximately £10 should be offered at each Wave.</li> </ul>	<ul style="list-style-type: none"> <li>• A marginal increase to £15 at Wave 4 is recommended.</li> </ul>
Parent interview	<ul style="list-style-type: none"> <li>• Interviews with the main parent are recommended for the first three Waves and interviews with second parents for at least Wave 1.</li> <li>• Second parent interviews at some or all waves could be conducted using the internet or post to reduce costs.</li> <li>• Testing the accuracy of collecting information by proxy for second parents could also lead to reduced costs.</li> </ul>	<ul style="list-style-type: none"> <li>• Interviews with both the main and second parent are recommended for at least the first three waves, but this will be dependent on the desired content of the study and the resources available.</li> </ul>

**Table 11.3 Recommendations for a second LSYPE cohort**

Topic	ESSENTIAL recommendations	DESIRABLE recommendations
<b>Data Linkage</b>	<ul style="list-style-type: none"> <li>It is recommended that data linkage with the National Pupil Database that currently occurs should continue.</li> <li>It is recommended that data linkage is also pursued with Health Episode Statistics and the Work and Pensions Longitudinal Study.</li> </ul>	<ul style="list-style-type: none"> <li>Other sources of administrative data could be considered, including criminal records held on the Police National Computer and Higher Education data.</li> </ul>
<b>Funding</b>	<ul style="list-style-type: none"> <li>Careful consideration should be given to the pitfalls of cross-departmental funding in regards to broadening the scope of the study.</li> <li>If cross-departmental funding is pursued, the distinctiveness and unique benefits of LSYPE need to be clearly defined and communicated to other Departments at the outset.</li> <li>Cross-sectional analysis of the first wave of the study should be encouraged, which would help to engage potential funding Departments early on in the life of the study.</li> <li>Ensure that the contributing Departments have opportunities to feed into the questionnaire design.</li> </ul>	
<b>Communication Plan</b>	<p><b>Communicating with users and respondents</b></p> <ul style="list-style-type: none"> <li>Traffic to study websites and responses to emails and SMS messages should be regularly monitored.</li> <li>Track who is accessing findings, in what format and how often in order to maximise interaction and improve communication.</li> </ul> <p><b>Communicating with respondents and those that affect them</b></p> <ul style="list-style-type: none"> <li>Prepare a communications plan including target audiences, branding and key messages, materials to be produced, timelines for distribution and measures for success.</li> </ul> <p><b>Communicating to study users</b></p> <ul style="list-style-type: none"> <li>Prepare a database of report/data recipients.</li> <li>Develop a single, professional looking and usable website.</li> <li>Prepare a distribution plan and timeline.</li> </ul>	<ul style="list-style-type: none"> <li>Appoint a communications manager to offer strategic advice about communications to young people and parents, develop materials, and ensure key messages are communicated consistently.</li> <li>Develop and test materials for parents and young people.</li> </ul>

### 11.3 For further discussion

As well as making recommendations for a second cohort, the review also highlighted areas that require further discussion before any decisions or recommendations can be made. Table 11.2 list these areas by topic.

**Table 11.4** Areas for further discussion

Chapter	Subject
Overview	<ul style="list-style-type: none"><li>Stakeholders agreed that the setting up of an Expert Advisory Panel for LSYPE would be useful. However, wider discussions are required to decide who would be best placed to act in an advisory capacity for this study in terms of the type of panel member (i.e. policy or analyst) and actual panel member (specific individuals).</li><li>The exact content of the thematic and policy relevant questions that could be answered using a second LSYPE need to be agreed in discussion with the DfE.</li></ul>
Sample & Fieldwork Design	<ul style="list-style-type: none"><li>The exact number and types of minority ethnic groups that should be boosted needs to be discussed further and in consideration to the funding available.</li><li>Final decisions on which stratification variables to use would have to be made following discussions of which subgroups are to be boosted, as these decisions are interdependent.</li></ul>
Mixed Modes	<ul style="list-style-type: none"><li>More empirical research is needed into generalised mode design. Development work is needed to identify whether particular questions can or cannot be asked in different modes.</li></ul>
Data Linkage	<ul style="list-style-type: none"><li>Prior to making requests for, or undertaking, data linkage, DfE need to consult with legal and ethical advisors to ensure that issues associated with confidentiality and informed consent are fully addressed.</li></ul>
Communications	<ul style="list-style-type: none"><li>The wording and mode of communication used to get messages about the study across to potential respondents should be developed and tested with young people before being rolled out. Messages for parents should also be developed and tested (these were not investigated as part of this review).</li></ul>

### 11.4 Final points

This review has shown just how unique and how successful LSYPE has been. Not only has the study generated large volumes of policy-relevant data for analysis but it is regularly used within Government Departments to inform policy development. The quality and availability of data has led to LSYPE being highly regarded by researchers, as evidenced by the fact that the data are used extensively in the academic community as well as by a wider range of users. While there are areas of overlap with other major cohort studies, LSYPE is the only contemporary major longitudinal study focusing on young people's experiences and entry into adulthood covering the whole of England. LSYPE can therefore offer unique insights into the

experiences of young people and the impact these may have on their later lives, and as such is a valuable resource for future and current policy.

A second LSYPE cohort would be invaluable in building on current knowledge of transitions from education in the teenage years through to early adulthood. It will not only serve to maximise the value of the first LSYPE cohort by enabling insightful comparative studies, but it will meet new current government priorities and provide economic value. The new government's focus on transparency can be served by second LSYPE. A new LSYPE cohort will ensure that there are no gaps in policy knowledge in the areas of school experiences, young people's relationships with their peers and families, their experiences of transitions into adulthood and work, their aspirations for the future and how these relate to their family circumstances and socioeconomic background. Other major cohort studies or multiple cross-sectional studies are unlikely to bridge these gaps. In addition, a single longitudinal study focusing solely on young people is likely to provide better value for money than commissioning a number of smaller cross-sectional studies.

# References

- Benabou, R. & Tirole, J. (2003) 'Intrinsic and extrinsic motivation' *Review of Economic Studies*, 70, pp. 489-520.
- Bhamra, S. Gatenby, R. Hacker, E. Killpack, C. Larkin, C., & Lessof, C. (2010) *Technical report of the National Child Development Study: 2008-2009 survey*. London: NatGen.
- Bradshaw, J. Bloor, K. Huby, M. Rhodes, D. Sinclair, I. Gibbs, I. Noble, M. McLennan, D. & Wilkinson, K. (2009) *Local Index of Child Well-Being: Summary Report*, London: Communities and Local Government.
- Burton, J. Laurie, H. & Lynn, P (2004) *The long term effectiveness of refusal conversion procedures on longitudinal surveys*, ISER Working Papers Number 2004-11.
- Calderwood, L. & Lessof, C. (2009) 'Enhancing Longitudinal Surveys by Linking to Administrative Data' in Lynn, P. (Ed.) *Methodology of Longitudinal Surveys*, Chichester, UK: Wiley.
- Christian, L. & Dillman, D. (2004) 'The Influence of Graphical and Symbolic Language Manipulations on Responses to Self-Administered Questions' *Public Opinion Quarterly*, 68, pp. 57-80.
- Christian, L. Dillman, D. A. & Smyth, J.D. (2008) 'The effects of mode and format on answers to scalar questions in telephone and web surveys', In, Lepkowski, J. Tucker, C. Brick, J. de Leeuw, E. Japac, L. Lavrakas, P. Link, W. & Sangster, R. (Eds.) *Advances in Telephone Survey Methodology*. New York: John Wiley Company.
- Church, A. (1993) 'Estimating the effect of incentives on mail survey response rates: a meta-analysis' *Public Opinion Quarterly*, 57, No.1, pp. 62-69.
- Collins, A. Bronte-Tinkew, J. & Burkhauser, M. (2008) *Using incentives to increase participation in out-of-school time programs*, Research-to-Results. [www.childtrends.org](http://www.childtrends.org).
- Deci, E. L. Koestner, R. & Ryan, R. M. (1999) 'A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation' *Psychological Bulletin*, 125, pp.627-668.
- de Leeuw, E. (2005) 'To mix or not to mix data collection modes in surveys', *Journal of Official Statistics*, 21, No. 2, pp. 233–255.

de Leeuw, E. (2008) 'Questionnaire design and measurement error', Presentation given on the 10<sup>th</sup> of April 2008 in London, Survey Measurement/ESRC Question Bank.

de Leeuw, E., Hox, J. & Dillman, D. (2008) 'Mixed mode surveys: when and why', *International Handbook of Survey Methodology*, Psychology Press.

Dillman, D. A. (1978) *Mail and telephone surveys: the total design method*, New York: Wiley.

Dillman, D. A. (2000) *Mail and internet surveys: the tailored design method*, New York: John Wiley.

Dillman, D. A. (2007) *Mail and internet surveys*, (2nd Edition), New Jersey: John Wiley.

Dillman, D.A. (2009) 'Some consequences of survey mode changes in longitudinal surveys', in Lynn, P. (Ed) *Methodology of longitudinal surveys*, Chichester: John Wiley & Sons Ltd, pp. 127-140.

Dillman, D. A. Brown, T. Carlson, J. Carpenter, E. Lorenz, F. Mason, R. Saltiel, J. & Sangster, R. (1995) 'Effects of category order on answers to mail and telephone surveys' *Rural Sociology*, 60, pp. 674–687.

Gray, J. C. Gatenby, R. & Huang, Y. (2010). *Millennium Cohort Study Sweep 4: Technical report*. London: NatCen.

Gouldner, A (1960) 'The norm of reciprocity: a preliminary statement' *American Sociological Review*, 25 (2), pp.161-178.

Goyder, J. (1994) 'An experiment with cash incentives on a personal interview survey', *Journal of Market Research Society*, 36 (4), pp. 360-366.

Groves, R. Cialdini & Couper, M. (1992) 'Understanding the decision to participate in a survey', *Public Opinion Quarterly*, 56, pp. 475-495.

Groves, R. & Couper, M. (1998) *Non-response in household interview surveys*, New York: Wiley.

Harbaugh, W. T. & Krause, K. (2000). 'Children's altruism in public good and dictator experiments,' *Artefactual Field Experiments 0046*, The Field Experiments Website.

Hawkes, D. & Plewis, I. (2006) 'Modelling non-response in the National Child Development Study, Journal of the Royal Statistical Society Series A' *Statistics in Society*, 169 part 3, pp.479-491.

Hochstim, J.R. (1967) 'A critical comparison of three strategies of collecting data from Households' *Journal of the American Statistical Association*. 62, pp. 976–989.

Holmberg, A. Lorenc, B. & Werner, P. (2008) *Optimal contact strategy in a mail and web mixed mode survey*, Proceedings of Q2008 European Conference on Quality in Official Statistics, Rome, Italy.

Hopkins, K. & Gullikson, A. (1992) 'Response rates in survey research: a meta-analysis of the effects of monetary gratuities', *Journal of Experimental Education*, 61(1), pp. 52-62.

Jackle, A. & Lynn, P. (2007) *Respondent incentives in a multi-mode longitudinal survey: cumulative effects on non-response and bias*, ISER Working Paper 2007-01, Colchester: University of Essex.

James, T. (1997) 'Results of Wave 1 incentive experiment in the 1996 Survey of Income and Program Participation. Proceedings of the Survey Research Section', *American Statistical Association*, pp. 834-839.

Japac, L. (1995) *Issues in Mixed mode Survey Design*, Stockholm: Statistics Sweden, March 1995.

Juster, F. T. & Smith, J. P. (1997) "Improving the Quality of Economic Data: Lessons from the HRS and AHEAD." *Journal of the American Statistical Association*, 92(440): pp 1268-1278.

Laurie, H. & Lynn, P. (2008) *The use of respondent incentives on longitudinal surveys*. ISER Working Paper 2008-42. Colchester: University of Essex.

Lengacher, J. E. Sullivan, C. M. Couper, M. P. & Groves, R. M. (1995) *Once reluctant, always reluctant? Effects of differential incentives on later survey participation in a longitudinal study*. Survey Research Center, University of Michigan.

Lynn, P. (2005) *Outcome categories and definitions of response rates for Panel surveys and other surveys involving multiple data collection events from the same units*, Unpublished manuscript. Colchester: University of Essex.

Lynn, P. & Sturgis, P. (1997) 'Boosting survey response through a monetary incentive and fieldwork procedures: an experiment', *Survey Methods Centre Newsletter*, 17(3), pp.18-22.

Lynn, P. Taylor, B. & Brook, L. (1997) 'Incentives, information and number of contacts: testing the effects of these factors on response to a longitudinal survey', *Survey Methods Newsletter*, 17(3), pp.13-17.

- Lynn, P. and K. Thomson, L. Brook (1998) 'An experiment with incentives on the British Social Attitudes Survey'. *Survey Methods Centre Newsletter*, 18(2): 12-14.
- Mack, S. Huggins, V. Keathley, D. & Sundukchi, M. (1998) 'Do monetary incentives improve response rates in the Survey of Income and Program Participation?' *Proceedings of the American Statistical Association, Survey Research Methods Section*: 529-534.
- Manfreda, K. L. Bosnjak, M. Berzelak, J. Hass, I. & Vehovar, V. (2008). Web surveys versus other survey modes: A meta-analysis comparing response rate. *International Journal of Market Research*, 50(1), 79-104.
- Martin, E. Abreu, D. & Winters, F. (2001) 'Money and motive: effects of incentives on longitudinal attrition in the Survey of Income and Program Participation', *Journal of Official Statistics*, 17 (2) pp. 267-284.
- Millar, M.M., O'Neill, A. C. & Dillman, D. A. (2009). Are Mode Preferences Real? Technical Report 09-003, Washington State University Social and Economic Sciences Research Center. Washington State University: Pullman.
- Pierce, W. D. & Cameron, J. (2002) 'A summary of the effects of reward contingencies on interest and performance' *The Behaviour Analyst Today*, 3 (2), pp. 221-228.
- Plewis, I. Smith, G. Wright, G. & Cullis, A. (2001) *Linking child poverty and child outcomes: Exploring data and research strategies*, London: Department for Work and Pensions.
- Rodgers, W. (2002) 'Size of incentive effects in a longitudinal study', *American Association for Public Research: Strengthening our Community – Section on Survey Research Methods*. Survey Research Center, University of Michigan.
- Salathiel, D. and Nicolaas, G. (2004) The Effects Of Monetary Incentives On Response Rates And Sample Bias. A paper presented at the International Conference on Transport Survey Quality and Innovation, Costa Rica, August 2004
- Shih, T., & Fan, X. (2008). Comparing response rates from Web and mail surveys: A meta-analysis. *Field Methods*, 20(3), 249-271.
- Singer, E. Grebler, N. Raghunathan, T. Van Hoewyk, J. & McGonagle, K. (1999) 'The effect of incentives in interviewer-mediated surveys', *Journal of Official Statistics* 15, pp.217-230.
- Singer, E. Groves, R. M. & Corning, A. D. (1999) 'Differential incentives: beliefs about practices, perceptions of equity, and effects on survey participation', *Public Opinion Quarterly*, 63 (2), pp. 251-260.

Singer, van Hoewyk and Couper, (1998) Unpublished analyses of data from the Survey of Consumer Attitudes during eight months in 1998 discussed in: Singer, E. and Kulka, RA. (2002) *Studies of Welfare Populations: Paying Respondents for Survey Participation*. Accessed on the web on 13 October 2010 at <http://aspe.hhs.gov/hsp/welf-res-data-issues02/04/04.htm>

Singer, E. & Kulka, R.A. (2000) *Paying respondents for survey participation*. Survey Methodology Program Working Paper No.092, Survey Research Center, Institute for Social Research, University of Michigan.

Stratford, N. Simmonds, N. & Nicolaas, G. (2003) National Travel Survey 2002: Incentives experiment report.  
[http://www.dft.gov.uk/stellent/groups/dft\\_transstats/documents/page/dft\\_transstats\\_023300.doc](http://www.dft.gov.uk/stellent/groups/dft_transstats/documents/page/dft_transstats_023300.doc)

Sutter, M. & Kocher, M. (2007) 'Trust and trustworthiness across different age groups', *Games and Economic Behavior*, 59 (2), pp. 364-382.

Thomas D. Frankenberg, E. & Smith, J. P. (2001) 'Lost but not forgotten: attrition and follow-up in the Indonesia Family Life Survey', *Journal of Human Resources*, 36, pp.556-592.

Voogt, R.J.J. & Saris, W.E. (2005) 'Mixed mode designs: Finding the balance between non-response bias and mode effects', *Journal of Official Statistics*, 21, pp. 367-87.

# Appendix A Advisory Panels of the first LSYPE cohort

It was initially proposed to establish an Academic Advisory Group of 30 to 40 academic experts in youth research for the first LSYPE cohort. Over 120 names emerged as potential advisors leading to a revision of the original model for the Advisory Group. It was clear that interests were focused in seven broad topic areas of relevance to the study and that a more effective mechanism for gaining advice would be to set up seven advisory panels.

Seven advisory panels were established:

- School policy and curriculum
- Ethnicity
- Adolescent development
- Family and health
- Labour market
- Youth policy and practise
- Methodology.

Convenors were appointed for each of these groups. The Advisory Panel members comprised people mainly from British Universities, but a number of other people were also recruited from European countries and from the USA.

The role of each group was to write a discussion paper identifying variables that the LSYPE needed to measure, and methods of measuring them. These papers were presented at a consultative conference held in 2002.

## **Family and Health**

Nicholas Emler - Advisory Panel Convenor (University of Surrey)

Glen Waller - Deputy Advisory Panel Convenor (St George's Hospital Medical School)

Kate Smith (Institute of Education), Amanda Sacker (University College London), John Hobcraft (London School of Economics), Robert Goodman (Institute of Psychiatry), Yvonne Kelly (University College London), Ingrid Schoon (City University), Kathleen Kiernan (London School of Economics), Mary Haines (Queen Mary College), Roger Ingham (University of Southampton), Russell Viner (Institute of Child Health, University College London), Sarah Hampson (University of Surrey), Tim Newburn (Goldsmiths College).

## **Labour Market**

Peter Dolton - Advisory Panel Convenor (London School of Economics)  
Abigail Mcknight - Deputy Advisory Panel Convenor (Institute for Fiscal Studies)  
Lorraine Dearden (Institute for Fiscal Studies), Lorna Unwin (University of Leicester), Walter Heinz (University of Bremen), Andy Furlong (University of Glasgow), David Ashton (Centre for Labour Market Studies, University of Leicester), Ken Roberts (University of Liverpool), Peter Elias (University of Warwick), Thomas Lange (North East Wales Institute of Higher Education), Anna Vignoles (London School of Economics), David Raffe (University of Edinburgh), Maria Iacovou (University of Essex), Stephen Machin (University College London), Richard Blundell (Institute for Fiscal Studies), Paul Gregg (University of Bristol), Gerry Makepeace (University of Cardiff), Walter Van Trier (University of Antwerp).

## **Methodology**

Ian Plewis - Advisory Panel Co-Convenor (Institute of Education)  
Peter Lynn - Advisory Panel Co-Convenor (University of Essex)  
Charlie Owen (Institute of Education), Howard Meltzer (Office of National Statistics), Geert Ridder (University of Southern California), Robert Michael (University of Chicago), Peter Dolton (Institute of Education), Dougal Hutchison (National Foundation for Educational Research), Keisuke Hirano (University of Miami), Harvey Goldstein (Institute of Education), Kevin Pickering (National Centre for Social Research), Peter Ratcliffe (University of Warwick), Vernon Gayle (University of Stirling).

## **School Policy and Curriculum**

David Gillborn - Advisory Panel Convenor (Institute of Education)  
Louise Morley - Deputy Advisory Panel Convenor (Institute of Education)  
Audrey Osler (University of Leicester), Inge Bates (University of Leeds), Bill Boyle (University of Manchester), Geoff Hayward (University of Oxford), Lesley Saunders (General Teaching Council).

## **Ethnicity**

Jagdish Gudara - Advisory Panel Convenor (Institute of Education)  
Charlie Owen - Deputy Advisory Panel Convenor (Institute of Education)  
Tozun Issa (Educational Consultant), Sonja Hall (Equality Action Team), Bill Boyle (University of Manchester), Nicola Rollock (The Runnymede Trust), Gajendra Verma (Birmingham Race Action Partnership), Richard Berthoud (University of Essex), Seamus Taylor (Commission for Racial Equality), John Singh (Education Consultant), Mark Blake (Winsor Fellowship), Joy Warmington (Birmingham Race Action Partnership), Ann Phoenix (Open University), Philomena Essed (University of Amsterdam), Heidi Mirza (Middlesex University), (Birmingham Race Action Partnership), Brian Richardson (Trades Union Congress), Ali Rattansi (City University), Jaqueline Eccles (University of Michigan), Jackie Beavan (Birmingham Race Action Partnership).

### **Youth Policy and Practice**

John Bynner - Advisory Panel Convenor (Institute of Education)

Bob Coles - Deputy Advisory Panel Convenor (University of York)

Manuela Bois-Reymond (Leiden University), Debi Roker (Trust for the Study of Adolescence), Gill Jones (Keele University), John Coleman (Trust for the Study of Adolescence), Tom Wylie (National Youth Agency), Ann Hagell (Policy Research Bureau), Peter Ratcliffe (University of Warwick), Howard Williamson (University of Cardiff), Tim Newburn (Goldsmiths College), Charlie Lloyd (Joseph Rowntree Foundation), Rob MacDonald (University of Teesside), Liza Catan (Trust for the Study of Adolescence).

### **Adolescent Development**

Nick Emler - Advisory Panel Convenor (University of Surrey)

Ingrid Schoon- Deputy Advisory Panel Convenor (City University)

Ann Hagell (Policy Research Bureau), Robert Goodman (Institute of Psychiatry), Hakan Stattin (Uppsala & Orebro Universities, Sweden), Rainer Silbereisen, Peter Noack (University of Jena), Sarah Hampson (University of Surrey), Alexander Grob (University of Bern), Debi Roker (Trust for the Study of Adolescence), Roger Ingham (University of Southampton), Anita Wilson (University of Lancaster), Jari-Erik Nurmi (University of Jyväskylä), Katariina Salmela-Aro (University of Jyväskylä), Mary Haines (Queen Mary Hospital), Helena Helve (University of Helsinki), Jenny Wickham (Institute of Psychiatry), Jacqueline Scott (University of Cambridge).

## Appendix B List of content of LSYPE (questionnaire)

Table B.1 Summary of questionnaire content, Waves 1 to 6

Respondent	Summary of content at Wave 1	Summary of content at Wave 2	Summary of content at Wave 3	Summary of content at Wave 4	Summary of content at Wave 5	Summary of content at Wave 6
<b>Household section</b>						
	Household Situation Household Grid Languages spoken in the home	Household Situation Household Grid Languages spoken in the home	Household Situation Household Grid Languages spoken in the home	Household Situation Household Grid Languages spoken in the home	Household Situation Household Grid <i>(section answered by YP at W5)</i>	Household Situation Household Grid <i>(section answered by YP at W6)</i>
<b>Main parent section</b>						
	Attitudes to the young person's school and involvement in education Extra Curricular classes Year 10 Subject choices - - - Special educational needs  Parental expectations and aspirations Family activities Household responsibilities and resources -	Attitudes to the young person's school and involvement in education - - School history Vocational courses Extra curricular classes Special educational needs  Parental expectations and aspirations - Household responsibilities and resources Young person history	Attitudes to the young person's school and involvement in education - - - - Extra curricular classes Special educational needs  Parental expectations and aspirations - Household responsibilities and resources -	Year 11 experiences ( <i>Boost only</i> ) * - - - - Extra curricular classes Special educational needs ( <i>Boost only</i> ) Post 16 plans ( <i>Boost only</i> ) ** - Household resources  Young person history (see	No parental interview at W5	No parental interview at W6

Table B.1 Summary of questionnaire content, Waves 1 to 6 continued ...

Respondent	Summary of content at Wave 1	Summary of content at Wave 2	Summary of content at Wave 3	Summary of content at Wave 4	Summary of content at Wave 5	Summary of content at Wave 6
	- Relationship with young person and contact with services Reasons for not living with natural parents Risk factors (absences, truancy, police contact, bullying)	Future contact details Relationship with young person and contact with services  Reasons for not living with natural parents Risk factors (absences, truancy, police contact, bullying)	Future contact details Relationship with young person and contact with services Reasons for not living with natural parents Risk factors (absences, truancy, police contact, bullying)	History section) Future contact details Relationship with young person and contact with services ( <i>Boost only</i> ) ***History section  ****Relationship with YP		
<b>Individual Parent section</b>						
	Demographics  Health Employment/activity history - -  Current activity  -	Demographics  Health - Employment/activity history since Wave one Employment/activity history for new entrants not interviewed at Wave one Current activity  Employment training and	- - - -  Current activity  -	Demographics ( <i>MP only</i> )  Health ( <i>MP and SP</i> ) - - -  Current economic activity  -	No parental interview at W5          [YP answered questions about parent(s) employment situation at W5]	No parental interview at W6

Table B.1 Summary of questionnaire content, Waves 1 to 6 continued ...

Respondent	Summary of content at Wave 1	Summary of content at Wave 2	Summary of content at Wave 3	Summary of content at Wave 4	Summary of content at Wave 5	Summary of content at Wave 6
	Qualifications and education - - -	earnings questions Qualifications and education  Benefits and tax credits - -	-  - Income estimate Second adult current economic activity ( <i>asked of MP only</i> )	Qualifications and education ( <i>MP only, but also answers about SP where applicable</i> ) - Income estimate ( <i>MP only</i> ) -  Job search ( <i>MP only</i> )		
<b>Young Person section</b>						
	Demographics - Attitudes to current school - - - Year 10 subject choices Rules and discipline - - - Homework	Demographics - - - - - Subjects studying and qualifications leading to - Reasons for Year 10 subject choices Homework	Demographics - - - - - Subjects being studied - - -	Demographics - Opinions on school Activities Jobs and training Activity history - - Qualifications being studied for now NEET - -	Demographics Attitudes on local area - Current activities Jobs and training Activity history - - Qualifications being studied for now NEET - -	Demographics  Current activities Jobs and training Activity history  Qualifications being studied for now NEET - -

Table B.1 Summary of questionnaire content, Waves 1 to 6 continued ...

Respondent	Summary of content at Wave 1	Summary of content at Wave 2	Summary of content at Wave 3	Summary of content at Wave 4	Summary of content at Wave 5	Summary of content at Wave 6
	ICT	ICT	-	-	-	-
	Study Support	Study Support	Study Support	-	-	-
	Future plans and advice	Future plans and advice	Future plans and advice	-	Information, advice and guidance	Information, advice and guidance
	-	-	-	Higher education plans	Higher education	Higher education students
	-	-	-	-	-	Potential higher education students
	Attitudes to school	Attitudes to school	Attitudes to school	-	-	-
	-	-	-	Attitudes to higher education and student debt	Attitudes to higher education and student debt	<i>(included in 'HE students' and 'Potential HE students')</i>
	Relations with parents	-	-	-	-	-
	-	Perceived discrimination	-	Perceived discrimination	-	-
	-	Knowledge of and intentions towards apprenticeships and related schemes	Knowledge of and intentions towards apprenticeships and related schemes	-	-	-
	-	Locus of control	-	-	-	-
	Household responsibilities	Household responsibilities	Household responsibilities	Caring responsibilities	Caring responsibilities	Childcare and caring responsibilities
	-	-	-	Looking after own children	Looking after own children	-
	-	-	-	-	Care to learn	Care to learn
	-	-	Education maintenance allowance (EMA)	EMA	EMA	-

Table B.1 Summary of questionnaire content, Waves 1 to 6 continued ...

Respondent	Summary of content at Wave 1	Summary of content at Wave 2	Summary of content at Wave 3	Summary of content at Wave 4	Summary of content at Wave 5	Summary of content at Wave 6
	- Use of leisure time  Risk factors (truancy, bullying, smoking, drugs) - - - - - - -	- Use of leisure time  Risk factors (truancy, bullying, smoking, drugs) General health over last few weeks (inc GHQ12) - - - - -	- -  Risk factors (truancy, bullying, smoking, drugs) Health  -  Parental Occupations  - - - -	- Spare time, car use and internet access Risk factors (truancy, bullying, smoking, drugs) General health, health and disability - - - - Data linkage consent	Job search - - - - Parental employment - Parent(s) employment Benefits Data linkage consent	- Sport (frequency)  Risk behaviours  Health and disability  Relationships and sexuality - - Income and benefits Data linkage consent
<b>History section</b>						
	Birth Health School history Choice of current school Sibling experience Relationship history -	Only if not asked at Wave 1 Only if not asked at Wave 1 Only if not asked at Wave 1 - Only asked if not asked at Wave 1 Relationship history -	- - - - - - -	Asked if boost Asked if boost - - Asked if boost  Asked if boost Living with Young Person	Not asked at wave 5 - - - - - -	Not asked at wave 6

*\*Questions were not exactly the same questions as in previous Waves \*\*Smaller set of questions than at Wave 3 \*\*\*Some questions in the Wave 4 history section ask about periods where the YP was not living with the MP but do not have the same variable names \*\*\*\*Relationship with YP section asks about police contact and absences (but not the same questions as at Wave 3)*

## Appendix C Publications on LSYPE

### LSYPE work published by DfE

Cebulla, A. & Tomaszewski, W. (2009) *Risky Behaviour and Social Activities*. Department for Children, Schools and Families, Research Report No DCSF-RR173. Available at:

<http://publications.dcsf.gov.uk/default.aspx?PageFunction=productdetails&PageMode=publications&ProductId=DCSF-RR173&>

Chowdry, H. & Crawford, C. and Goodman, A. (2009) *Drivers and barriers to educational success - evidence from the longitudinal study of young people in England*. Department for Children, Schools and Families, Research Report No DCSF-RR102

Available at:

<http://publications.dcsf.gov.uk/default.aspx?PageFunction=productdetails&PageMode=publications&ProductId=DCSF-RR102&>

DCSF Statistical Bulletin (2008) *YCS & LSYPE: the Activities & Experiences of 16 year olds: England 2007*.

Available at:

<http://www.dcsf.gov.uk/rsgateway/DB/SBU/b000795/index.shtml>

DCSF Statistical Bulletin (2009) *YCS & LSYPE: the Activities & Experiences of 16 year olds: England 2008*.

Available at:

<http://www.dcsf.gov.uk/rsgateway/DB/SBU/b000850/index.shtml>

DCSF Statistical Bulletin (2010) *YCS & LSYPE: the Activities & Experiences of 16 year olds: England 2009*.

Available at:

<http://www.dcsf.gov.uk/rsgateway/DB/SBU/b000937/index.shtml>

Demack, S. Platts-Fowler, D. Robinson D. Stevens, A. & Wilson, I. (2010) *Young People and Community Cohesion: Analysis from the Longitudinal Study of Young People in England (LSYPE)*. Department for Education, Research Report No DFE-RR033

Available at:

<http://publications.education.gov.uk/eOrderingDownload/DFE-RR033.pdf>

Green, R. & Ross, A (2010) *Young people's alcohol consumption and its relationship to other outcomes and behaviour*, Department for Education, Research Report No DFE-RR005.

Available at: <http://www.education.gov.uk/research/data/uploadfiles/DFE-RR005.pdf>

Green, R. Collingwood, A. & Ross, A. (2010) *The Characteristics of Bullying Victims in Schools*. Department for Education, Research Report No DFE-RR001.

Available at: <http://www.education.gov.uk/research/data/uploadfiles/DFE-RR001.pdf>

Piess, A & Kalton Westat, G. (2009) *A Strategy for Handling Missing Data in the Longitudinal Study of Young People in England (LSYPE)*. Department for Children, Schools and Families, Research Report No DCSF-RW086.

Available at:

<http://publications.dcsf.gov.uk/default.aspx?PageFunction=productdetails&PageMode=publications&ProductId=DCSF-RW086&>

Ross, A. (2009) *Disengagement from Education among 14-16 year olds*, Department for Children, Schools and Families, Research Report No DCSF-RR178.

Available at:

<http://publications.dcsf.gov.uk/default.aspx?PageFunction=productdetails&PageMode=publications&ProductId=DCSF-RR178&>

Strand, S. (2007) *Minority Ethnic Pupils in the Longitudinal Study of Young People in England (LSYPE)*. Department for Children, Schools and Families, Research Report No DCSF-RR002.

Available at:

<http://publications.dcsf.gov.uk/default.aspx?PageFunction=productdetails&PageMode=publications&ProductId=DCSF-RR002&>

Strand, S. (2008) *Minority Ethnic Pupils in the Longitudinal Study of Young People in England: Extension Report on Performance in Public Examinations at Age 16*.

Department for Children, Schools and Families, Research Report No DCSF-RR029.

Available at:

<http://publications.dcsf.gov.uk/default.aspx?PageFunction=productdetails&PageMode=publications&ProductId=DCSF-RR029&>

## **PUBLICATIONS BY OTHER GOVERNMENT DEPARTMENTS**

### **Cabinet Office**

Cabinet Office (2008) *Aspirations and attainment among young people in deprived*.

For further information see:

[http://www.cabinetoffice.gov.uk/media/109339/aspirations\\_evidence\\_pack.pdf](http://www.cabinetoffice.gov.uk/media/109339/aspirations_evidence_pack.pdf)

### **BIS**

Bates, P. Pollard, E. Usher, T. Oakley, J. (2009) *Who is Heading for HE? Young People's Perceptions of, and Decisions About, Higher Education (2009)*.

For further information:

[http://www.employment-studies.co.uk/pubs/report.php?id=bis\\_rp\\_003](http://www.employment-studies.co.uk/pubs/report.php?id=bis_rp_003)

## **Defra/DfE**

Dickerson, A. & McIntosh, S. (2009) *Post – 16 Educational Choices in Rural Areas A Report Prepared for the Department for Environment, Food and Rural Affairs by the Department of Economics, University of Sheffield.* (Published jointly with DfE).

Available at:

<http://publications.dcsf.gov.uk/default.aspx?PageFunction=productdetails&PageMode=publications&ProductId=DEFRA-2009-12-01+&>

Green, R. (2009) *Educational Attainment in Rural Areas, A Report Prepared for the Department for Environment, Food and Rural Affairs and the Department for Education by the National Centre for Social Research.* Available at:

<http://publications.dcsf.gov.uk/default.aspx?PageFunction=productdetails&PageMode=publications&ProductId=DEFRA-2009-12-02+&>

## **DWP**

Platt, L. (2009) *Ethnicity and Child Poverty Feasibility Study.* Working Paper No 69, Department for Work and Pensions. External Analysis.

## **PUBLISHED BY NON-GOVERNMENT DEPARTMENTS**

Bradley, S. Migali, G. & Taylor, J. (2010) *Funding, school specialisation and test scores: An evaluation of the specialist schools policy using matching models.*

Lancaster University Management School Working Paper 2010/007

Available at:

<http://www.lums.lancs.ac.uk/publications/viewpdf/006416/>

Bradley, S. Migali, G. & Taylor, J. (2009) *An evaluation of the impact of funding and school specialisation on student performance using matching models.*

Lancaster University Management School Working Paper 2009/006

Available at:

<http://www.lums.lancs.ac.uk/publications/viewpdf/005894/>

De Coulon, A. Meschi, E. Swaffield, J. Vignoles, A. & Wadsworth, J. (2010) *Minimum Wage and Staying-on Rates in Education for Teenagers.* Research report for the Low Pay Commission.

Available at:

[http://www.lowpay.gov.uk/lowpay/research/pdf/Decoulon\\_et\\_al\\_final\\_07\\_jan\\_10.pdf](http://www.lowpay.gov.uk/lowpay/research/pdf/Decoulon_et_al_final_07_jan_10.pdf)

Dickerson, A. & McIntosh, S. (2010) *The impact of distance to nearest education institution on the post-compulsory education participation decision.* Working Paper, Department of Economics, University of Sheffield.

Available at:

<http://www.shef.ac.uk/economics/research/serps/year.html>

Foreman-Peck, J. & Foreman-Peck, L. (2006) *Cognitive Achievement and Behaviour in School*. Welsh Institute for Research in Economics and Development, Cardiff Business School & School of Education, University of Northampton.

Available at:

<http://www.bristol.ac.uk/cmipo/plug/workshops/peck.pdf>

Fumagalli, E. & Fumagalli, L. (2010) *Like oil and water or chocolate and peanut butter? Ethnic diversity and social participation of young people in England*.

University of Padua Research Project

Available at:

<http://www.cepr.org/meets/wkcn/2/2413/papers/Fumagalli.pdf>

Gregg, P. & Macmillan, L. (2009) *Family Income and Education in the Next Generation: Exploring income gradients in education for current cohorts of youth*.

CMPO and Department of Economics, University of Bristol, CMPO Working Paper Series No. 09/223.

Available at:

<http://www.bristol.ac.uk/cmipo/publications/papers/2009/wp223.pdf>

Harris, A. Staff, J. & Sabates, R. (2008) *Oppositional culture beyond the United States? Perceptions of Discrimination and Academic Engagement in the United Kingdom*.

Available at:

<http://www.princeton.edu/~angelh/Website/Ogbu%20Studies/Harris,%20Staff,%20Sabates%20-%20ASR.pdf>

Howieson, C. Croxford, C. & Howat, N. (2008) *Meeting the Needs for Longitudinal Data on Youth Transitions in Scotland - An Options Appraisal*. A joint report to the Scottish Government Schools Directorate and Lifelong Learning Directorate.

Available at:

<http://www.scotland.gov.uk/Publications/2008/10/17104150/0>

Irwin, S. (2009) 'Family contexts, norms and young people's orientations: researching *diversity*', *Journal of Youth Studies*, 12, Issue 4, pp. 337 – 354.

Available at:

<http://prod.informaworld.com/smpp/content~content=a912281127~db=all~jumptype=rss>

Lynn, P. (2009) *Methodology of Longitudinal Surveys*. John Wiley & Sons Ltd.

Available at:

<http://onlinelibrary.wiley.com/doi/10.1002/9780470743874.ch1/summary>

Sartarelli, M. (2009) *Do achievement labels affect students' well-being? Evidence from discontinuities in test scores in England*. Institute of Education, University of London.

Available at:

[http://www.eale.nl/Conference2009/Programme/PapersB/add102789\\_wu7FJuzdxo.pdf](http://www.eale.nl/Conference2009/Programme/PapersB/add102789_wu7FJuzdxo.pdf)

Silva, O. & Gibbons, S. (2009) *School Quality, Child Wellbeing and Parents' Satisfaction*. Centre for the Economics of Education. DP103.

Available at:

<http://cee.lse.ac.uk/cee%20dps/ceedp103.pdf>

Skaliotis, E. (2009) 'Changes in parental involvement in secondary education: an exploration study using the Longitudinal Study of Young People in England', *British Educational Research Journal*

Strand, S. (in press) 'The limits of social class in explaining ethnic gaps in educational attainment', *British Educational Research Journal*, University of Warwick.

Available at:

<http://wrap.warwick.ac.uk/2363/>

Sullivan, A. Zimdars, A. & Heath, A. (2010) 'The social structure of the 14-16 curriculum in England', *International Studies in Sociology of Education*, 20, Issue 1, pp. 5 – 21.

### **Course work material**

LSYPE data are also used as study material within a number of university programmes, for example as part of the coursework for a Market Research Society Diploma. MRS is an awarding body for qualifications in market and social research. One of the units in their Level 7 Diploma qualification is 'Analysing and Interpreting Quantitative Market & Social Research Data'; for this unit, a dataset is selected from the UK Data Archive to form the basis of the candidate assessment. An example of the questions from a current paper which has been set by the MRS for one of the Diploma modules is provided below:

### **Task 1: The Research Objectives & Sampling Approaches**

This task requires you to analyse and evaluate the research design and sampling methods used for this research study, from the drawing of the original sample to the current Wave 5 sampling and boost samples.

In your answer you should:

- identify the objectives of this study – what problems/issues do the project set out to address; giving example outcomes for each objective
- identify the approaches taken to sampling. Discuss the strengths and weaknesses of each of these approaches in relation to the given research objectives, and with reference to any relevant statistical theory
- evaluate the appropriateness of the sample design, size and composition of the quantitative samples with reference to the given research objectives and analysis plan
- discuss the implications of this study being longitudinal and based on a longitudinal versus an independent sampling approach. Identify strategies that have been used to maintain the longitudinal study (including incentives) and any further strategies that could be used to increase the response rate
- discuss the over-sampling used in selecting schools at stage 1 of the sample design, and the sample boosts applied at stage 2 of the sample design for selecting pupils. Discuss the impact of these sampling decisions on analysis data from the LSYPE.

# Appendix D Content and review of other datasets

## **Growing Up in Scotland**

Growing Up in Scotland is a large scale longitudinal social survey designed to examine the characteristics, circumstances and behaviours of children from birth to late adolescence (and possibly beyond). It forms a central part of the Scottish Government's strategy for the long term monitoring and evaluation of its policies for children, with a specific focus on the early years. Although the survey has various features in common with other cohort projects, such as the Millennium Cohort Study, it also differs in a number of important respects. For example, it has a specifically and uniquely Scottish focus. It is driven specifically by the needs of policy, it has a particular focus on service use, awareness and contact in key stages of childhood - e.g. health, education, childcare and it has a greater focus on the early years of children's lives. Although the study addresses a wealth of policy information the context is different than LSYPE due to it being Scotland-only.

The main topics covered are childcare, education, social work, health and social inclusion.

### Household

- Household membership and relationships in the household
- Neighbourhood
- Housing and accommodation.

### History

- Pregnancy and birth
- Infant feeding
- Child development
- Early experiences of pre-school
- Early experiences of primary school.

### Young person

- Health record linkage
- School record linkage
- Child health
- Height and weight
- Activities
- Food and eating
- Childcare
- Child social networks
- Cognitive assessments.

## Parent

- Partner's interview
- Non-resident parents
- Parental socio-demographics
- Parenting support
- Parenting styles, activities and attitudes
- Parental relationships and responsibilities
- Parent and family social networks
- Parental health
- Employment
- Income and financial stress.

## **The Families and Children Study (FACS)**

FACS (formerly known as the Survey of Low Income Families (SOLIF)) is a longitudinal refreshed Panel survey which began in 1999. The tenth and final Wave of FACS was carried out in 2008. Substantial changes were made to the survey in the third Wave, when the name was changed: From 2001 (Wave 3) onwards the focus has been on *all* families irrespective of income. This provides data on Government targets towards reducing child poverty.

The focus of FACS is on younger children and the sample of children is much smaller than LSYPE. There is less focus on children's own experiences and more on the experiences of their parents.

The main aims of FACS are to analyse the effect of work incentives, and measure the effects of policy on families' living standards. Additionally, FACS explores changes in family circumstances over time, monitors the impact of benefits & tax credits (Family Tax Credit and Child Tax Credit) in supporting families with young children, explores barriers to work, particularly for low income families, and the measures to overcome such barriers, as well as explores more general family welfare issues. FACS 2003, 2004, 2006 and 2008 (Waves 5, 6, 8 and 10) included a 10 minute self completion questionnaire for all children aged 11-15 in the family.

The main interview includes the following core topics: household structure, health of the respondent and children, caring activities, household and living arrangements, education and training, work and working in previous two years, experience of income support, benefits received, savings and income, living standards, job searching, prior relationships, attitudes and morale. The shorter partner interview has the following topics: health, education, work, training, job search.

## Household:

- Information about the family unit
- Family composition
- Housing
- Social Capital.

Young person:

- Health
- School and education
- Problems and use of local services
- Childcare arrangements.

Main parent:

- Relationship history
- Contact with non-resident partners
- Receipt of benefits and tax credits
- Income and savings
- Expenditure and hardship
- Education and training
- Health
- Caring responsibilities
- Employment and self-employment
- Work history
- Unemployment and job search
- Parental aspirations for children.

Partner interview:

- Education and training
- Health
- Employment and self-employment
- Earnings
- Unemployment and job search
- Caring responsibilities.

### **Understanding Society**

Understanding Society is a major new research study designed to provide valuable evidence about the people of the UK, their lives, experiences, attitudes, health, behaviours and beliefs. Understanding Society's fieldwork began in 2009 and it aims to track 100,000 people living in Britain. The survey questions cover a wide spectrum of areas relating to respondents' working and personal lives. It will tell us about: their state of health, experiences of crime, personal finances, bringing up children, how involved people are in their local community, employment, and views and outlook, including views on the political system. The broad aim is to understand the life journey that people take.

Unlike LSYPE, the main focus is not on young people. It is designed to provide information on major social trends rather than specifically the development of young people so would benefit from comparisons alongside LSYPE in terms of analysis.

Understanding Society will collect a wide range of information on a great number of topics over time, including:

- Standard of living measures (income, consumption, material deprivation, expenditure, financial well-being).
- Family, social networks and interactions, local contexts, social support, technology and social contacts.
- Attitudes and behaviours related to environmental issues (energy, transport, air quality, global warming etc).
- Illicit and risky behaviour (crime, drug use, anti-social behaviour etc).
- Lifestyle, social, political, religious and other participation, identity and related practices, dimensions of life satisfaction/happiness.
- Psychological attributes cognitive abilities and behaviour.
- Preferences, beliefs, attitudes and expectations.
- Health outcomes and health related behaviour.
- Education, human capital and work.
- Initial conditions, life history.

### **Millennium Cohort Study**

MCS is a multi-disciplinary research project following the lives of around 19,000 children born in the UK in 2000/2001. It is the most recent of Britain's world-renowned national longitudinal birth cohort studies. MCS's field of enquiry covers such diverse topics as parenting; childcare; school choice; child behaviour and cognitive development; child and parental health; parents' employment and education; income and poverty; housing, neighbourhood and residential mobility; and social capital and ethnicity. As MCS is strong on socio-economic data, a second LSYPE focusing more on other areas such as attitudes and aspirations and life transitions would complement this study. As these two cohorts of young people would be almost the same age in 2012 (MCS respondents aged 12, LSYPE respondents aged 13/14), a second LSYPE cohort would need to be clearly differentiated from MCS. LSYPE has a strong focus on school factors, which could be expanded upon in a new cohort to further differentiate the study from other cohorts of similar age. Another important advantage of a second LSYPE is that pupils would be clustered by school, enabling researchers to consider school effectiveness, which the MCS does not allow.

MCS comprises a main interview (usually with the mother), a partner interview (if there is a partner in the household), a paper child-self completion questionnaire, child physical measurements, child cognitive assessments, a follow-up teacher survey, and various data linkage consents. The interview with the mother collects information on the household, the mother themselves and the child. The interview with the partner collects information mainly on the partner themselves. The areas covered by these elements are as follows:

Main interview:

- Household membership and relationships
- Languages spoken at home
- Family context
- Schooling & Childcare
- Child Activities
- Parenting Activities
- Child Health
- Parent's Health
- Employment, Income & Education
- Housing and Local Area

Partner interview:

- Parenting Activities
- Parent's Health
- Employment, Education, Income

Child:

- Height, weight, body fat percentage and waist measurement
- Cognitive assessments (varied with age). At age seven these were Sally and Anne, Word reading, Progress in Maths and Pattern Construction.
- Child self-completion questionnaire, covering the following topics:
  - Hobbies
  - Their relationship with their friends
  - Their feelings e.g. how often they feel happy or sad
  - What they do at school

Teacher:

- Child's abilities
- Child's behaviour
- Child profile and special needs
- Parent attitudes
- The child's class
- The teacher themselves

**Table D.1 Summary of main aims and topics covered by other key datasets**

Study	Main Aims	Main Topics Covered	Distinct Modules	Overlap with LSYPE
<ul style="list-style-type: none"> <li>GUS</li> </ul>	<p>Examine characteristics, circumstances and behaviours of children from birth to late adolescence.</p> <p>Monitor and evaluate Scottish policies for children (specific focus on early years).</p>	<ul style="list-style-type: none"> <li>Childcare</li> <li>Education</li> <li>Social work</li> <li>Health</li> <li>Social inclusion</li> </ul>	<ul style="list-style-type: none"> <li>Interview with main carer</li> <li>Separate interview with partner in second Wave</li> </ul>	<ul style="list-style-type: none"> <li>Education</li> <li>Parental views</li> </ul>
<ul style="list-style-type: none"> <li>FACS</li> </ul>	<p>Analyse the effect of work incentive measures.</p> <p>Measure the effects of policy on families' living standards.</p> <p>Explore changes in family circumstances over time.</p> <p>Monitor the impact of benefits and tax credits in supporting families with young children.</p> <p>Explore barriers to work, especially for low income families.</p> <p>Explore general family welfare issues.</p>	<ul style="list-style-type: none"> <li>Household structure</li> <li>Health of respondent &amp; children</li> <li>Caring activities</li> <li>Household &amp; living arrangements</li> <li>Education &amp; training</li> <li>Work</li> <li>Experience of income support</li> <li>Benefits, savings &amp; income</li> <li>Living standards</li> <li>Job searching</li> <li>Prior relationships</li> <li>Attitudes &amp; morale</li> </ul>	<ul style="list-style-type: none"> <li>Interview with main respondent (usually mother)</li> <li>Short interview with partner</li> <li>Self completion questionnaire (Waves 1-4)</li> <li>Self completion questionnaire for children aged 11-15 (Waves 5, 6 and 8)</li> </ul>	<ul style="list-style-type: none"> <li>Household structure</li> <li>Living circumstances and social position</li> <li>Self completion for young people (younger than LSYPE)</li> </ul>
<ul style="list-style-type: none"> <li>MCS</li> </ul>	<p>Collect information on the child's formative years with a view to seeing how these 'initial conditions' impact on their later life.</p>	<ul style="list-style-type: none"> <li>Household membership and relationships</li> <li>Languages spoken at home</li> <li>Family context</li> <li>Schooling &amp; Childcare</li> <li>Child Activities</li> <li>Parenting Activities</li> <li>Child Health</li> <li>Parent's Health</li> <li>Employment, Income &amp; Education</li> <li>Housing and Local Area</li> </ul>	<ul style="list-style-type: none"> <li>Interview with main respondent (usually mother)</li> <li>Interview with partner</li> <li>Paper self completion questionnaire for cohort child (at age 7 only)</li> <li>Child physical measurements</li> <li>Child cognitive assessments</li> <li>Follow-up teacher survey</li> </ul>	<ul style="list-style-type: none"> <li>Household structure</li> <li>Living circumstances and social economic position</li> <li>Parents employment and education</li> <li>Self completion for young people (younger than LSYPE)</li> </ul>
<ul style="list-style-type: none"> <li>BCS70</li> </ul>	<p>Examine social and biological characteristics of mothers in relation to neonatal morbidity.</p> <p>Produce results comparable to previous birth cohorts.</p>	<ul style="list-style-type: none"> <li>Biological characteristics</li> <li>Physical development</li> <li>Educational development</li> <li>Social development</li> <li>Economic development</li> </ul>	<ul style="list-style-type: none"> <li>Parental interviews</li> <li>Medical tests</li> <li>Aptitude tests</li> <li>Child interviews from age 10</li> </ul>	<ul style="list-style-type: none"> <li>Parental information</li> <li>Economic data</li> <li>Education</li> </ul>

**Table D.1 Summary of main aims and topics covered by other key datasets continued...**

<ul style="list-style-type: none"> <li>EPPSE 3-14</li> </ul>	<p>Explore the way the influences of school interact with social and family background in shaping attainment, attitudes and social/behavioural adjustment over time.</p>	<ul style="list-style-type: none"> <li>School influences</li> <li>Child and family influences</li> <li>Neighbourhood influences</li> <li>Out of school learning</li> <li>Resilient and vulnerable pupils</li> </ul>	<ul style="list-style-type: none"> <li>Young person interviews</li> <li>Parental interviews</li> </ul>	<ul style="list-style-type: none"> <li>School information</li> <li>Family relationships</li> <li>Vulnerable pupils</li> </ul>
<ul style="list-style-type: none"> <li>ALSPAC</li> </ul>	<p>Study genetic and environmental factors contributing to long term health and development.</p>	<ul style="list-style-type: none"> <li>Health and medical history of mother and grandparents</li> <li>Psychological wellbeing of parents</li> <li>Parental attitudes and behaviours</li> <li>Parents' childhood experiences</li> <li>Parental social support</li> <li>Socio-economic information</li> <li>Child health and lifestyle</li> <li>Child development</li> <li>School information</li> </ul>	<ul style="list-style-type: none"> <li>Interview with mother</li> <li>Child self completion questionnaire</li> <li>Psychical measurements</li> <li>Psychological tests</li> <li>Extraction of DNA</li> </ul>	<ul style="list-style-type: none"> <li>Parental attitudes</li> <li>School information</li> </ul>
<ul style="list-style-type: none"> <li>US</li> </ul>	<p>Understand life journeys people take.</p> <p>Catch major trends in society.</p> <p>Gain understanding of why major changes in the way we live and work take place.</p>	<ul style="list-style-type: none"> <li>Standard of living</li> <li>Family, social networks and interactions</li> <li>Attitudes and behaviours related to environmental issues</li> <li>Illicit and risky behaviour</li> <li>Social participation</li> <li>Psychological attributes</li> <li>Beliefs and attitudes</li> <li>Health</li> <li>Education and work</li> <li>Life history</li> </ul>	<ul style="list-style-type: none"> <li>Household questionnaire</li> <li>Interviews with all adults</li> <li>Proxy interviews for those not present</li> <li>Adult self completion questionnaire</li> <li>Youth self completion questionnaire (ages 10-15)</li> <li>Blood samples</li> </ul>	<ul style="list-style-type: none"> <li>Family relationships</li> <li>Self completion questionnaires</li> <li>Risky behaviours</li> <li>Attitudes</li> <li>Education</li> </ul>

## Appendix E Youth policies that may affect the analysis questions

- Over 3,500 Parent Support Advisers and similar professionals are currently helping parents across England to be fully involved in their children's learning and are working with schools and families to support better pupil behaviour and attendance. There have also been significant changes to the KS3 curriculum which are projected over a three year time scale, and became statutory for Year 7 pupils in 2008. From September 2009 the new curriculum was applied to all year 7 and 8 pupils, and from September 2010 it will apply across years 7, 8 and 9. Greater flexibility has been introduced, in order to engage and motivate students and to encourage them to see the links between academic subjects. The key three aims are to 1) enable the students to see the relevance of the whole curriculum, 2) teach the skills that students need to succeed, and 3) give students 'real-life' experiences to give them an opportunity to use the skills learned at school.
- The National Challenge, introduced in 2008, has been implemented to challenge and support secondary schools with low GCSE results in order to help them improve. The project ensures that head teachers are supported by advisors and can draw upon a range of options (including carefully targeted help for teaching and learning) to develop strong leadership and flexibility in the design of local solutions. By 2011, the National Challenge target is to ensure that at least 30% of KS4 pupils in schools achieve at least 5 good GCSEs including English and Maths. There is also the 'Making Good Progress' Pilot which has been introduced in order to offer one-to-one tuition for those pupils who need it most, specifically targeted to English and Maths. Tuition will be offered to 10% of pupils in KS 2 and 3 (up to year 9).
- In 2009, in conjunction with the DH, the former DCSF launched the 'Child Health Strategy: Healthy Lives, Brighter Futures' which 1) provides additional support during early years and strengthens the role of Sure Start Centres, 2) develops the 'Healthy Child Programme' for school children and their families, and 3) provides opportunities for young people including access to sport and young people friendly health services. The DfE have introduced food standards which mean that school lunches must now include at least 1 portion of fruit and 1 portion of vegetables contributing to the '5 A Day' initiative. Limits have been placed on the amount of fat and saturated fat that can be included in a meal and schools can no longer provide snacks and drinks which have no nutritional value. DfE have also launched a healthy cookbook 'Real Meals – Simple Cooking Made Easy' free to all Year 7 pupils to help them learn healthy versions of old favourites.
- In 2008 the former DCSF published the 'Youth Alcohol Action Plan' (along with DH & HO) to set out how the government will tackle the issue of teenage

alcohol consumption through the following 5 objectives, 1) stepping up enforcement activity to address young people drinking in public places, 2) taking action with industry on young people and alcohol, 3) developing a national consensus on young people and drinking, 4) establishing a new partnership with parents on teenage drinking and 5) supporting young people to make sensible decisions about alcohol.

- The 'Your Child, Your Schools, Our Future' White Paper explained the changing role of schools at the centre of a community working together to support all children. Schools now offer extended services based on the needs of their local community, which include; study support, before and after school activities, childcare, parenting and family support, access to specialist health and social care services, and community facilities such as adult and family learning.
- The former DCSF introduced 'Family Intervention Projects' to provide help and support for those families that need it.
- The 'Youth Crime Action Plan' was published in 2008 and outlined what can be done to tackle the immediate and long term causes of youth crime. The government have 1) introduced street based teams to deter people from becoming involved in crime, 2) expanded 'Operation Staysafe' which uses existing child protection legislation to remove children and young people from the street late at night, 3) increased after school police patrols to tackle anti-social behaviour, 4) provided Youth Offending Team Workers in police custody suites so that there is action to tackle young people's offending at the point of arrest, 5) established more Family Intervention Projects across England, and 6) expanded 'Safer Schools Partnerships'.

## Appendix F Cohort studies with Respondent Advisory Panels

### **English Longitudinal Study of Ageing (held in 2005):**

This study originally invited respondents to take part in focus groups by means of a Christmas card in 2004, but no responses were received. Instead, respondents were randomly selected from a sub-sample living in Leeds and London (40 respondents from each) to receive a letter, which stated that *if they did not wish to take part* in the focus groups they should respond. Respondents who did not opt out were then telephoned and asked to participate. No incentives were offered to respondents.

All focus group respondents were core sample members interviewed at both Waves of the study that had been completed to date. Those excluded included those who had had a proxy interview or partial interview at Wave 2, those who did not want to be contacted, and those who were living in institutions. The Leeds respondents were all selected to be 65 years or over, whereas the London respondents were all under the age of 65. The respondents were also sampled on the basis of gender, whether they had concurrent or single interviews, and if they were working or retired (in the younger group only).

Topics explored included the respondents' understanding of the study, their motivations for being involved, their positive and negative experiences of involvement in the study and how these experiences could be improved. This included questionnaire content and ways of keeping in touch (Christmas cards, newsletters with study results, and fridge magnets with contact numbers).

### **Families and Children Study (held in 2002):**

Letters of invitation were sent to all the potential respondents, offering those who did not want to take part the chance to withdraw. Screening interviews were then conducted by telephone and those selected were invited to participate. All those who participated were given £15 as a gesture of thanks.

Respondents were selected by purposive sampling and allocated to five groups: respondents interviewed in consecutive Waves who were still interested in taking part; respondents interviewed in previous Waves who gave permission to be re-contacted but subsequently refused at Wave 3; respondents who had previously refused but were then interviewed at Wave 3; first-time interviewees at Wave 3 who agreed to re-contact; and, first-time interviewees at Wave 3 who did not agree to re-contact. Quotas of respondents were filled according to age, family type (lone parent or two parents), number of children, age of youngest child and location.

Topics explored were the respondents' views on ways of keeping in touch, the information sent to respondents and how this was perceived, strategies for obtaining permission to re-call, contact with interviewers, interview length and content and the incentives used.

**Communities Study (held in 2008):**

This exercise recruited respondents who had not previously been involved in a NatCen survey and compared these with ex-respondents from the Communities Study. Respondents who had not previously responded to a survey were recruited using a recruitment agency who were given a screening questionnaire and instructed to fill quotas. Respondents who had already taken part in the Communities Study and who had indicated that they were happy to be re-contacted were also sampled using a screening questionnaire. Characteristics selected on were gender and age: Respondents who had previously taken part in the Communities Study were recruited from the 30-55 age group, and those who had not were recruited from the 18-30 and 55+ age groups, with an equal gender balance in each. No incentive was given to take part in the focus groups.

Respondents were asked for their responses to different advance letters asking them to participate in surveys, including whether incentives (vouchers or stamps) should be included with these letters. There was no discussion of the actual survey content in these focus groups.

## Appendix G Response Rates

### **Responsive design**

According to work published by Groves & Heeringa (2006), field operations can be more effectively monitored and managed to deliver higher quality data for the same cost. This technique, called responsive design, monitors key statistics during the data collection process and chooses optimal times to introduce protocol changes. An optimal time is when it is no longer cost effective to continue applying effort under the existing protocol. By introducing a protocol change at this time, both cost and bias concerns are addressed. First, sample persons with desirable characteristics, who were not attracted to the prior protocol, may be drawn into the respondent pool by the new protocol. This better balances the data and therefore reduces non-response bias. Second, cost is saved by introducing the change once the impact of the prior protocol has been maximized, thereby efficiently cutting off the supply of resources applied to a protocol resulting in minimal returns on response rate and no reduction on bias. Additional cost savings can be incorporated by disproportionately selecting cases which are expected to have a high propensity of responding, according to non-response models, and by delegating fieldwork assignments in clustered areas.

The types of protocol changes that could be incorporated into a responsive design are varied. The technique relies on paradata (auxiliary information) to capture characteristics of the cases which is used for identification of desirable characteristics to address bias or to develop propensity models. One example, pertinent to the LSYPE, is the identification of the teenager's sex from administrative records at the initial Wave. This allows for targeted and tailored follow-up with either boys or girls after the initial protocol has been maximized, in order to balance the sample. The tailoring may involve a specific type of contact or incentive determined to be appealing to the demographic. The follow-up is typically more targeted and costly than the protocol used in the data collection up to that point and therefore resources are preserved until they are needed.

Responsive design is typically applied on cross-sectional studies or at the first Wave of a longitudinal study but the technique could be applied to later Waves. The process would work in a similar fashion with data from prior Waves becoming the all-important paradata allowing for the identification of cases necessary to balance the respondent pool as well as more predictive propensity modelling. If there is no sub-selection of follow-up cases and all outstanding cases are pursued, this prior Wave information could be used to tailor interventions for different kinds of non-respondents or even at the outset of the Wave. For example, mode preferences could be collected at the first Wave and subsequent Waves could deliver the survey in the preferred mode.

# Response Trees

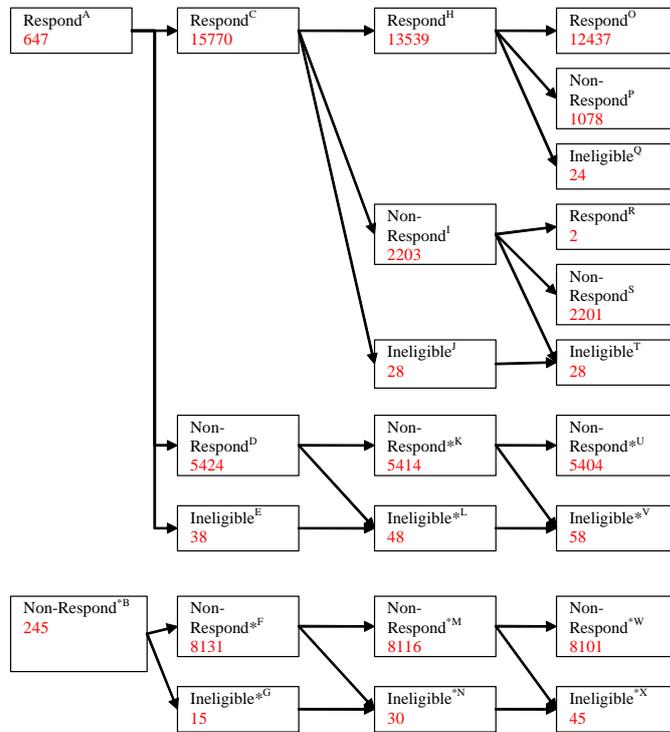
## Stage 1

## Stage 2

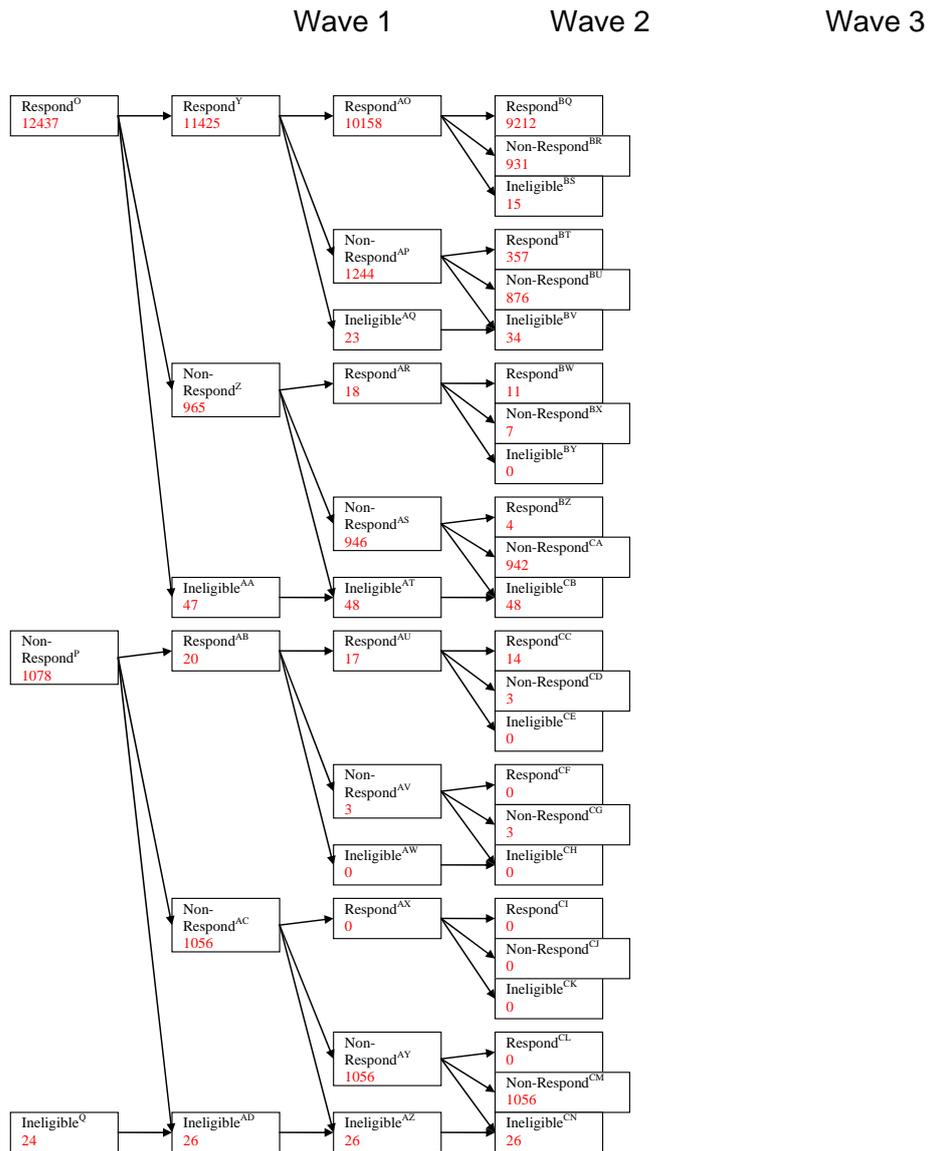
Wave 1

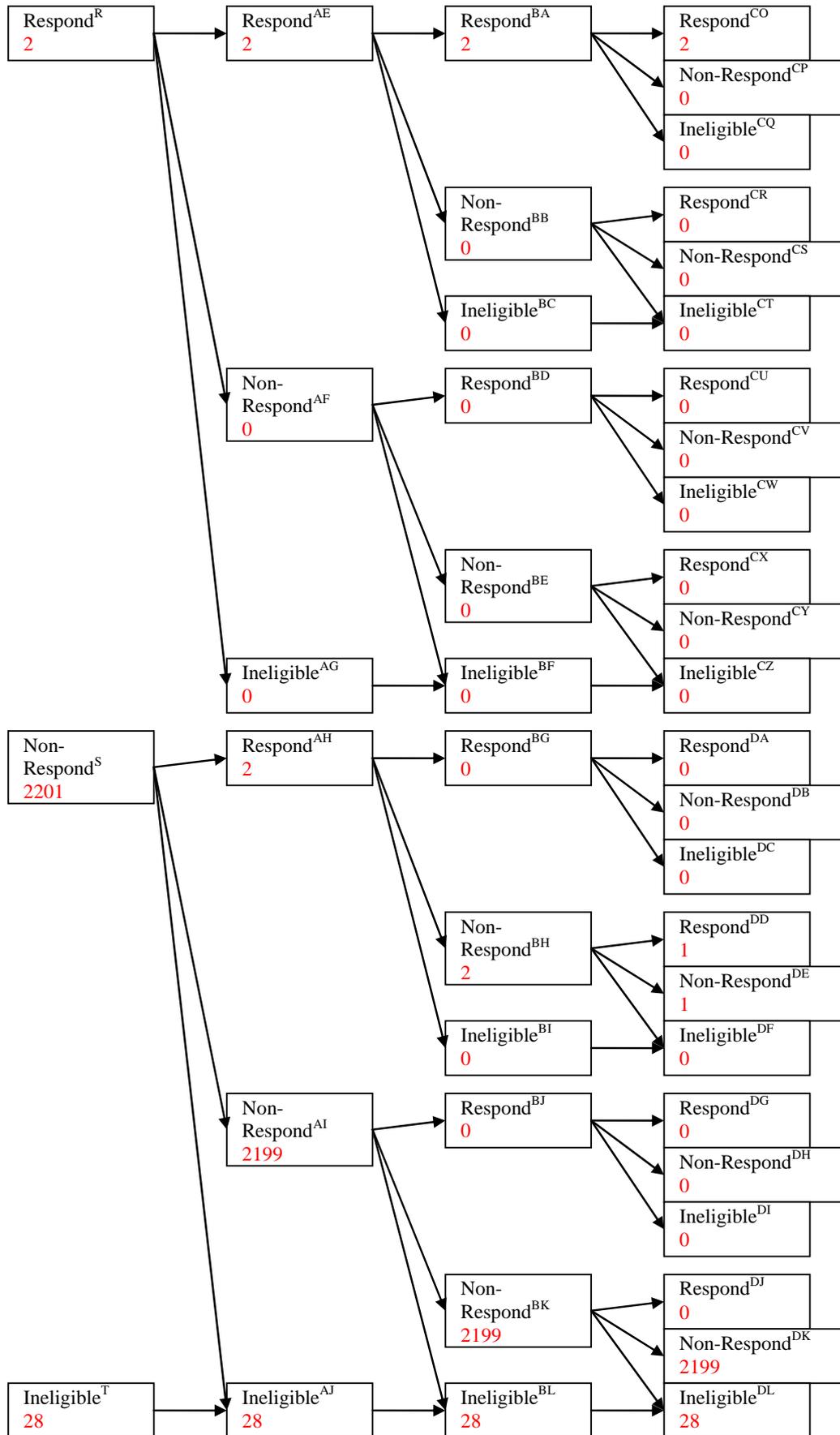
Wave 2

Wave 3



## Stage 2





## Calculation of estimates in the response trees

The percentage of issued cases in Wave 1 that were ineligible =  $E/(C+D+E) = 38/(15770+5424+38) = 38/21232 = 0.18\%$

$$G^* = 8146 * 0.18\% = 15$$

The percentage of respondents in Wave 1 who became ineligible at Wave 2 =  $J/C = 28/15770 = 0.18\%$

$$L^* = 38 + 5424 * 0.18\% = 38 + 10 = 48$$

$$N^* = 15 + 8131 * 0.18\% = 15 + 15 = 30$$

The percentage of respondents in Wave 2 who became ineligible at Wave 3 =  $Q/H = 24/13539 = 0.18\%$

$$V^* = 5414 * 0.18\% + 48 = 10 + 48 = 58$$

$$X^* = 8116 * 0.18\% + 30 = 15 + 30 = 45$$

The percentage of respondents in Wave 3 who became ineligible at Wave 4 =  $(AA + AG) / (O + R) = (47+0)/(12437+2) = 47/12439 = 0.38\%$

$$AL^* = 5404 * 0.38\% + 58 = 21 + 58 = 79$$

$$AN^* = 8101 * 0.38\% + 45 = 31 + 45 = 76$$

Percentage of those who responded in Wave 4 who became ineligible at Wave 5 =  $(AQ+AW+BC+BI)/(Y+AB+AE+AH) = (23+0+0+0)/(11425+20+2+2) = 23/11449 = 0.2\%$

$$BN^* = 5383 * 0.2\% + 79 = 11 + 79 = 90$$

$$BP^* = 8070 * 0.2\% + 76 = 16 + 76 = 92$$

Percentage of those who responded in Wave 5 who became ineligible in Wave 6 =  $(BS+BY+CE+CK+CQ+CW+DC+DI)/(AO+AR+AU+AX+BA+BD+BG+BJ) = 15/10195 = 0.15\%$

$$DN^* = 5372 * 0.15\% + 90 = 8 + 90 = 98$$

$$DP^* = 8054 * 0.15\% + 92 = 12 + 92 = 104$$

## Calculation of response rates

### ***Cross-sectional Unconditional (including schools that didn't cooperate and non-responding pupils in Wave 1)***

#### *Wave 1*

$$\begin{aligned}RR_1 &= (\text{respond in Wave 1})/(\text{eligible sample members in Wave 1}) \\ &= C/(C+D+F^*) \\ &= 15770/(15770+5424+8131) \\ &= 53.8\%\end{aligned}$$

#### *Wave 2*

$$\begin{aligned}RR_2 &= (\text{respond in Wave 2})/(\text{eligible sample members in Wave 2}) \\ &= H/(H+I+K^*+M^*) \\ &= 13539/(13539+2203+5414+8116) \\ &= 46.3\%\end{aligned}$$

#### *Wave 3*

$$\begin{aligned}RR_3 &= (\text{respond in Wave 3})/(\text{eligible sample members in Wave 3}) \\ &= (O+R)/(O+R+P+S+U^*+W^*) \\ &= (12437+2)/(12437+2+1078+2201+5404+8101) \\ &= 42.6\%\end{aligned}$$

#### *Wave 4*

$$\begin{aligned}RR_4 &= (\text{respond in Wave 4})/(\text{eligible sample members in Wave 4}) \\ &= (Y+AB+AE+AH)/(Y+Z+AB+AC+AE+AF+AH+AI+AK^*+AM^*) \\ &= (11425+20+2+2)/(11425+965+20+1056+2+0+2+2199+5383+8070) \\ &= 39.3\%\end{aligned}$$

Wave 5

$$\begin{aligned} RR_5 &= (\text{respond in Wave 5})/(\text{eligible sample members in Wave 5}) = \\ & (AO+AR+AU+AX+BA+BD+BG+BJ)/(AO+AP+AR+AS+AU+AV+AX+AY+BA+BB+BD+ \\ & BE+BG+BH+BJ+BK+BM^*+BO^*) = \\ & (10158+18+17+0+2+0+0+0)/(10158+1244+18+946+17+3+0+1056+2+0+0+0+0+2+0 \\ & +2199+5372+8054) = 35.1\% \end{aligned}$$

Wave 6

$$\begin{aligned} RR_6 &= (\text{respond in Wave 6})/(\text{eligible sample members in Wave 6}) = \\ & (BQ+BT+BW+BZ+CC+CF+CI+CL+CO+CR+CU+CX+DA+DD+DG+DJ)/(BQ+BR+BT \\ & +BU+BW+BX+BZ+CA+CC+CD+CF+CG+CI+CJ+CL+CM+CO+CP+CR+CS+CU+CV \\ & +CX+CY+DA+DB+DD+DE+DG+DH+DJ+DK+DM^*+DO^*) = \\ & (9212+357+11+4+14+0+0+0+2+0+0+0+0+1+0+0)/(9212+931+357+876+11+7+4+94 \\ & 2+14+3+0+3+0+0+0+1056+2+0+0+0+0+0+0+0+0+0+0+1+1+0+0+0+2199+5364+8042) \\ & = 33.1\% \end{aligned}$$

***Cross-sectional Unconditional (excluding schools that didn't cooperate but including non-responding pupils in Wave 1)***

Wave 1

$$\begin{aligned} RR_1^a &= (\text{respond in Wave 1})/(\text{eligible sample members in Wave 1 in cooperating schools}) \\ & = C/(C+D) \\ & = 15770/(15770+5424) \\ & = 74.4\% \end{aligned}$$

Wave 2

$$\begin{aligned} RR_2^a &= (\text{respond in Wave 2})/(\text{eligible sample members in Wave 2 in cooperating schools}) \\ & = H/(H+I+K^*) \\ & = 13539/(13539+2203+5414) \\ & = 64.0\% \end{aligned}$$

Wave 3

$RR_3^a$  = (respond in Wave 3)/(eligible sample members in Wave 3 in cooperating schools)

$$= (O+R)/(O+R+P+S+U^*)$$

$$= (12437+2)/(12437+2+1078+2201+5404)$$

$$= 58.9\%$$

#### Wave 4

$RR_4^a$  = (respond in Wave 4)/(eligible sample members in Wave 4 in cooperating schools)

$$= (Y+AB+AE+AH)/(Y+Z+AB+AC+AE+AF+AH+AI+AK^*)$$

$$= (11425+20+2+2)/(11425+965+20+1056+2+0+2+2199+5383)$$

$$= 54.4\%$$

#### Wave 5

$RR_5^a$  = (respond in Wave 5)/(eligible sample members in Wave 5 in cooperating schools) =

$$(AO+AR+AU+AX+BA+BD+BG+BJ)/(AO+AP+AR+AS+AU+AV+AX+AY+BA+BB+BD+BE+BG+BH+BJ+BK+BM^*) =$$

$$(10158+18+17+0+2+0+0+0)/(10158+1244+18+946+17+3+0+1056+2+0+0+0+0+2+0+2199+5372) = 48.5\%$$

#### Wave 6

$RR_6^a$  = (respond in Wave 6)/(eligible sample members in Wave 6 in cooperating schools) =

$$(BQ+BT+BW+BZ+CC+CF+CI+CL+CO+CR+CU+CX+DA+DD+DG+DJ)/(BQ+BR+BT+BU+BW+BX+BZ+CA+CC+CD+CF+CG+CI+CJ+CL+CM+CO+CP+CR+CS+CU+CV+CX+CY+DA+DB+DD+DE+DG+DH+DJ+DK+DM^*) =$$

$$(9212+357+11+4+14+0+0+0+2+0+0+0+0+1+0+0)/(9212+931+357+876+11+7+4+942+14+3+0+3+0+0+0+1056+2+0+0+0+0+0+0+0+0+0+0+1+1+0+0+0+2199+5364) = 45.8\%$$

### ***Cross-sectional Conditional on having responded at Wave 1 (excluding schools that didn't cooperate)***

#### Wave 2

$RR_2^b$  = (respond in Wave 2)/(eligible sample members in Wave 2 in cooperating schools, excluding non-responding pupils in Wave 1)

$$= H/(H+I)$$

$$= 13539/(13539+2203)$$

$$= 86.0\%$$

### Wave 3

$RR_3^b$  = (respond in Wave 3)/(eligible sample members in Wave 3 in cooperating schools, excluding non-responding pupils in Wave 1)

$$= (O+R)/(O+R+P+S)$$

$$= (12437+2)/(12437+2+1078+2201)$$

$$= 79.1\%$$

### Wave 4

$RR_4^b$  = (respond in Wave 4)/(eligible sample members in Wave 4 in cooperating schools, excluding non-responding pupils in Wave 1)

$$= (Y+AB+AE+AH)/(Y+Z+AB+AC+AE+AF+AH+AI)$$

$$= (11425+20+2+2)/(11425+965+20+1056+2+0+2+2199)$$

$$= 73.1\%$$

### Wave 5

$RR_5^b$  = (respond in Wave 5)/(eligible sample members in Wave 5 in cooperating schools, excluding non-responding pupils in Wave 1) =

$$(AO+AR+AU+AX+BA+BD+BG+BJ)/(AO+AP+AR+AS+AU+AV+AX+AY+BA+BB+BD+BE+BG+BH+BJ+BK) =$$

$$(10158+18+17+0+2+0+0+0)/(10158+1244+18+946+17+3+0+1056+2+0+0+0+0+2+0+2199) = 65.2\%$$

### Wave 6

$RR_6^b$  = (respond in Wave 6)/(eligible sample members in Wave 6 in cooperating schools, excluding non-responding pupils in Wave 1) =

$$(BQ+BT+BW+BZ+CC+CF+CI+CL+CO+CR+CU+CX+DA+DD+DG+DJ)/(BQ+BR+BT+BU+BW+BX+BZ+CA+CC+CD+CF+CG+CI+CJ+CL+CM+CO+CP+CR+CS+CU+CV+CX+CY+DA+DB+DD+DE+DG+DH+DJ+DK) =$$

$$(9212+357+11+4+14+0+0+0+2+0+0+0+0+1+0+0)/(9212+931+357+876+11+7+4+942+14+3+0+3+0+0+0+1056+2+0+0+0+0+0+0+0+0+0+1+1+0+0+0+2199) = 61.5\%$$

**Cross-sectional Conditional (conditional on having responded at the previous Wave)**

*Wave 2*

$RR_2^c$  = (respond in Wave 1 and Wave 2)/(eligible sample members in Wave 2 – if also respond in Wave 1)

$$= H/(H+I)$$

$$= 13539/(13539+2203)$$

$$= 86.0\%$$

*Wave 3*

$RR_3^c$  = (respond in Wave 2 and Wave 3)/(eligible sample members in Wave 3 – if also respond in Wave 2)

$$= O/(O+P)$$

$$= (12437)/(12437+1078)$$

$$= 92.0\%$$

*Wave 4*

$RR_4^c$  = (respond in Wave 3 and Wave 4)/(eligible sample members in Wave 4 – if also respond in Wave 3)

$$= (Y+AE)/(Y+Z+AE+AF)$$

$$= (11425+2)/(11425+965+2+0)$$

$$= 92.2\%$$

*Wave 5*

$RR_5^c$  = (respond in Wave 4 and Wave 5)/(eligible sample members in Wave 5 – if also respond in Wave 4)

$$= (AO+AU+BA+BG)/(AO+AP+AU+AV+BA+BB+BG+BH)$$

$$= (10158+17+2+0)/(10158+1244+17+3+2+0+0+2)$$

$$= 89.1\%$$

#### Wave 6

$$\begin{aligned} RR_6^c &= (\text{respond in Wave 5 and Wave 6})/(\text{eligible sample members in Wave 6 – if also respond in Wave 5}) = \\ &= (BQ+BW+CC+CI+CO+CU+DA+DG)/(BQ+BR+BW+BX+CC+CD+CI+CJ+CO+CP+CU \\ &+CV+DA+DB+DG+DH) = \\ &= (9212+11+14+0+2+0+0+0)/(9212+931+11+7+14+3+0+0+2+0+0+0+0+0+0) = \\ &= 90.8\% \end{aligned}$$

#### ***Longitudinal Unconditional (including schools that didn't cooperate and non-responding pupils in Wave 1)***

#### Wave 2

$$\begin{aligned} RR_2^d &= (\text{respond in Waves 1 and 2})/(\text{eligible sample members in Waves 1 and 2}) \\ &= H/(H+I+K^*+M^*) \\ &= 13539/(13539+2203+5414+8116) \\ &= 46.3\% \end{aligned}$$

#### Wave 3

$$\begin{aligned} RR_3^d &= (\text{respond in Waves 1, 2 and 3})/(\text{eligible sample members in Waves 1, 2 and 3}) \\ &= O/(O+P+R+S+U^*+W^*) \\ &= 12437/(12437+1078+2+2201+5404+8101) \\ &= 42.6\% \end{aligned}$$

#### Wave 4

$$\begin{aligned} RR_4^d &= (\text{respond in Waves 1,2,3 and 4})/(\text{eligible sample members in Waves 1,2,3 and 4}) \\ &= Y/(Y+Z+AB+AC+AE+AF+AH+AI+AK^*+AM^*) \\ &= 11425/(11425+965+20+1056+2+0+2+2199+5383+8070) \\ &= 39.2\% \end{aligned}$$

#### Wave 5

$$\begin{aligned} RR_5^d &= (\text{respond in Waves 1,2,3,4 and 5})/(\text{eligible sample members in Waves 1,2,3,4 and 5}) = \\ & AO/(AO+AP+AR+AS+AU+AV+AX+AY+BA+BB+BD+BE+BG+BH+BJ+BK+BM^*+BO^*) \\ &= 10158/(10158+1244+18+946+17+3+0+1056+2+0+0+0+0+2+0+2199+5372+8054) \\ &= 34.9\% \end{aligned}$$

#### Wave 6

$$\begin{aligned} RR_6^d &= (\text{respond in Waves 1,2,3,4,5 and 6})/(\text{eligible sample members in Waves 1,2,3,4,5 and 6}) = \\ & BQ/(BQ+BR+BT+BU+BW+BX+BZ+CA+CC+CD+CF+CG+CI+CJ+CL+CM+CO+CP+ \\ & CR+CS+CU+CV+CX+CY+DA+DB+DD+DE+DG+DH+DJ+DK+DM^*+DO^*) = \\ & 9212/(9212+931+357+876+11+7+4+942+14+3+0+3+0+0+0+1056+2+0+0+0+0+0+0 \\ & +0+0+0+1+1+0+0+0+0+2199+5364+8042) = 31.7\% \end{aligned}$$

#### ***Longitudinal Unconditional (excluding schools that didn't cooperate but including non-responding pupils in Wave 1)***

#### Wave 2

$$\begin{aligned} RR_2^e &= (\text{respond in Waves 1 and 2})/(\text{eligible sample members in Waves 1 and 2 in cooperating schools}) \\ &= H/(H+I+K^*) \\ &= 13539/(13539+2203+5414) \\ &= 64.0\% \end{aligned}$$

#### Wave 3

$$\begin{aligned} RR_3^e &= (\text{respond in Waves 1,2 and 3})/(\text{eligible sample members in Waves 1,2 and 3 in cooperating schools}) \\ &= O/(O+R+P+S+U^*) \\ &= 12437/(12437+2+1078+2201+5404) \\ &= 58.9\% \end{aligned}$$

#### Wave 4

$$RR_4^e = (\text{respond in Waves 1,2,3 and 4})/(\text{eligible sample members in Waves 1,2,3 and 4 in cooperating schools})$$

$$\begin{aligned}
&= Y/(Y+Z+AB+AC+AE+AF+AH+AI+AK^*) \\
&= 11425/(11425+965+20+1056+2+0+2+2199+5383) \\
&= 54.3\%
\end{aligned}$$

#### Wave 5

$$\begin{aligned}
RR_5^e &= (\text{respond in Waves 1,2,3,4 and 5})/(\text{eligible sample members in Waves 1,2,3,4 and 5 in cooperating schools}) \\
&= \\
&AO/(AO+AP+AR+AS+AU+AV+AX+AY+BA+BB+BD+BE+BG+BH+BJ+BK+BM^*) \\
&= 10158/(10158+1244+18+946+17+3+0+1056+2+0+0+0+0+2+0+2199+5372) \\
&= 48.3\%
\end{aligned}$$

#### Wave 6

$$\begin{aligned}
RR_6^e &= (\text{respond in Waves 1,2,3,4,5 and 6})/(\text{eligible sample members in Waves 1,2,3,4,5 and 6 in cooperating schools}) = \\
&BQ/(BQ+BR+BT+BU+BW+BX+BZ+CA+CC+CD+CF+CG+CI+CJ+CL+CM+CO+CP+ \\
&CR+CS+CU+CV+CX+CY+DA+DB+DD+DE+DG+DH+DJ+DK+DM^*) = \\
&9212/(9212+931+357+876+11+7+4+942+14+3+0+3+0+0+0+1056+2+0+0+0+0+0+0 \\
&+0+0+0+1+1+0+0+0+2199+5364) = 43.9\%
\end{aligned}$$

#### ***Longitudinal Conditional on having responded at Wave 1 (excluding schools that didn't cooperate)***

#### Wave 2

$$\begin{aligned}
RR_2^f &= (\text{respond in Waves 1 and 2})/(\text{eligible sample members in Waves 1 and 2 in cooperating schools, excluding non-responding pupils in Wave 1}) \\
&= H/(H+I) \\
&= 13539/(13539+2203) \\
&= 86.0\%
\end{aligned}$$

#### Wave 3

$$RR_3^f = (\text{respond in Waves 1,2 and 3})/(\text{eligible sample members in Waves 1,2 and 3 in cooperating schools, excluding non-responding pupils in Wave 1})$$

$$= O/(O+R+P+S)$$

$$= 12437/(12437+2+1078+2201)$$

$$= 79.1\%$$

#### Wave 4

$$RR_4^f = (\text{respond in Waves 1,2,3 and 4})/(\text{eligible sample members in Waves 1,2,3 and 4 in cooperating schools, excluding non-responding pupils in Wave 1})$$

$$= Y/(Y+Z+AB+AC+AE+AF+AH+AI)$$

$$= 11425/(11425+965+20+1056+2+0+2+2199)$$

$$= 72.9\%$$

#### Wave 5

$$RR_5^f = (\text{respond in Waves 1,2,3,4 and 5})/(\text{eligible sample members in Waves 1,2,3,4 and 5 in cooperating schools, excluding non-responding pupils in Wave 1})$$

$$= AO/(AO+AP+AR+AS+AU+AV+AX+AY+BA+BB+BD+BE+BG+BH+BJ+BK)$$

$$= 10158/(10158+1244+18+946+17+3+0+1056+2+0+0+0+0+2+0+2199)$$

$$= 64.9\%$$

#### Wave 6

$$RR_6^f = (\text{respond in Waves 1,2,3,4,5 and 6})/(\text{eligible sample members in Waves 1,2,3,4,5 and 6 in cooperating schools, excluding non-responding pupils in Wave 1}) =$$

$$BQ/(BQ+BR+BT+BU+BW+BX+BZ+CA+CC+CD+CF+CG+CI+CJ+CL+CM+CO+CP+CR+CS+CU+CV+CX+CY+DA+DB+DD+DE+DG+DH+DJ+DK) =$$

$$9212/(9212+931+357+876+11+7+4+942+14+3+0+3+0+0+0+1056+2+0+0+0+0+0+0+0+0+0+1+1+0+0+0+2199) = 59.0\%$$

### ***Longitudinal Conditional (conditional on having responded at the previous Wave)***

#### Wave 2

$$RR_2^g = (\text{respond in Waves 1 and 2})/(\text{eligible sample members in Waves 1 and 2 – if also respond in Wave 1})$$

$$= H/(H+I)$$

$$= 13539/(13539+2203)$$

$$= 86.0\%$$

#### Wave 3

$RR_3^g$  = (respond in Waves 1,2 and 3)/(eligible sample members in Waves 1,2 and 3 – if also respond in Wave 2)

$$= O/(O+P)$$

$$= (12437)/(12437+1078)$$

$$= 92.0\%$$

#### Wave 4

$RR_4^g$  = (respond in Waves 1,2,3 and 4)/(eligible sample members in Waves 1,2,3 and 4 – if also respond in Wave 3)

$$= Y/(Y+Z+AE+AF)$$

$$= 11425/(11425+965+2+0)$$

$$= 92.2\%$$

#### Wave 5

$RR_5^g$  = (respond in Waves 1,2,3,4 and 5)/(eligible sample members in Waves 1,2,3,4 and 5 – if also respond in Wave 4)

$$= AO/(AO+AP+AU+AV+BA+BB+BG+BH)$$

$$= 10158/(10158+1244+17+3+2+0+0+2)$$

$$= 88.9\%$$

#### Wave 6

$RR_6^c$  = (respond in Waves 1,2,3,4,5 and 6)/(eligible sample members in Waves 1,2,3,4,5 and 6 – if also respond in Wave 5) =  
 $BQ/(BQ+BR+BW+BX+CC+CD+CI+CJ+CO+CP+CU+CV+DA+DB+DG+DH) =$   
 $9212/(9212+931+11+7+14+3+0+0+2+0+0+0+0+0+0) = 90.5\%$

**Table G.1 Respondents versus non-respondents, Wave 2**

Wave 2	Productive %	Refusal %	Non-contact/ mover %	Other unproductive %
Owner occupied	73.2	66.3	52.0	52.0
Household does not have use of a motor vehicle	13.2	16.5	23.8	22.1
English is not the main/first language of the household	6.0	6.9	13.0	6.5
Lone parent household	22.6	23.6	39.4	38.1
MP has non-white ethnicity	11.5	13.7	22.4	16.1
Urban	80.0	83.1	85.8	84.8
Independent school	7.2	7.2	8.9	5.3
MP is male	14.3	16.9	16.9	13.6
Number of dependent children in the household (mean)	2.2	2.3	2.4	2.3
MPs age (mean)	42.1	41.7	41.1	40.5
<i>Main parents employment status</i>				
Employed	73.4	63.5	59.3	60.9
Unemployed	1.6	2.4	3.2	4.1
Economically inactive	24.9	34.1	37.5	34.9
<i>Family's NSSEC</i>				
Higher and lower managerial and professional occupations	40.1	33.5	30.3	29.7
Intermediate occupations	7.2	6.8	9.2	7.3
Small employers and own account workers	12.2	15.9	11.1	12.4
Lower supervisory and technical occupations	11.9	9.5	9.5	13.6
Semi-routine/ routine occupations	24.3	26.8	27.8	29.4
Never worked/ Long term unemployed	4.2	7.4	12.1	7.5
<i>Main parents highest educational qualification</i>				
Degree or equivalent or higher education below degree	25.9	19.2	22.2	15.2
GCE A Level or equivalent GCSE grades A_C or equivalent	44.4	41.8	35.0	47.8
Qualifications at level 1 or below, or other qualifications	10.9	13.0	12.4	12.2
No qualifications	18.8	26.0	30.5	24.9
<i>Main parents marital status</i>				
Single (never married)	6.1	7.5	11.9	14.6
Cohabiting (married or partner)	76.0	75.0	58.7	59.9
Separated, divorced, widowed or other	17.9	17.5	29.4	25.5

**Table G.1 Respondents versus non-respondents, Wave 2 continued ...**

Wave 2	Productive %	Refusal %	Non-contact/ mover %	Other unproductive %
<i>Gross household income</i>				
Less than £10,400	15.9	23.0	27.7	26.0
£10,400 less than £16640	17.5	17.5	22.0	23.0
£16640 less than £26000	21.5	22.2	19.2	23.4
£26000 less than £36400	19.2	14.4	14.2	11.4
£36400 or more	25.9	22.9	16.9	16.1
<i>Government office region</i>				
North East	5.3	4.5	4.6	1.7
North West	15.3	16.5	16.4	20.7
Yorkshire and The Humber	10.6	9.1	9.7	9.2
East Midlands	8.4	6.9	6.5	8.8
West Midlands	11.5	11.7	8.7	14.0
East of England	10.6	10.4	12.3	5.0
London	12.4	16.2	21.8	15.3
South East	16.4	16.0	13.3	16.7
South West	9.5	8.7	6.7	8.7
<i>Deprivation index</i>				
1 Least deprived	25.2	21.7	17.6	16.2
2	23.7	21.9	16.4	17.7
3	21.2	21.6	20.9	24.6
4	16.5	18.4	26.2	20.4
5 Most deprived	13.3	16.5	18.9	21.1
<i>Base unweighted</i>	<i>13415</i>	<i>888</i>	<i>977</i>	<i>275</i>
<i>Base weighted</i>	<i>13535</i>	<i>871</i>	<i>911</i>	<i>253</i>

**Table G.2 Respondents versus non-respondents, Wave 3**

<b>Wave 3</b>	<b>Productive %</b>	<b>Refusal %</b>	<b>Non-contact/ mover %</b>	<b>Other unproductive %</b>
Owner occupied	75.0	68.2	40.1	53.9
Household does not have use of a motor vehicle	12.5	16.3	28.7	25.1
English is not the main/first language of the household	5.8	6.4	6.8	7.6
Lone parent household	21.3	22.9	37.9	40.6
MP has non-white ethnicity	11.2	15.4	14.4	16.8
Urban	79.5	81.7	90.4	88.5
Independent school	7.4	5.3	4.3	8.7
MP is male	13.8	15.5	13.6	16.7
Number of dependent children in the household (mean)	2.1	2.1	2.3	2.0
MPs age (mean)	43.4	42.4	41.2	41.6
<i>Main parents employment status</i>				
Employed	74.1	71.7	54.4	66.7
Unemployed	1.6	1.2	3.4	2.9
Economically inactive	24.2	27.1	42.2	30.4
<i>Family's NSSEC</i>				
Higher and lower managerial and professional occupations	42.6	32.5	22.5	29.1
Intermediate occupations	7.0	8.1	9.1	8.4
Small employers and own account workers	8.9	9.1	6.1	10.1
Lower supervisory and technical occupations	13.0	14.6	11.2	10.7
Semi-routine/ routine occupations	24.5	30.1	40.2	30.3
Never worked/ long term unemployed	4.0	5.6	11.0	11.5
<i>Main parents highest educational qualification</i>				
Degree or equivalent or higher education below degree	27.4	17.8	19.9	23.3
GCE A Level or equivalent GCSE grades A_C or equivalent	44.5	44.1	39.0	39.5
Qualifications at level 1 or below, or other qualifications	11.6	14.1	13.2	12.5
No qualifications	16.5	24.1	27.9	24.7
<i>Main parents marital status</i>				
Single (never married)	8.6	11.0	19.5	21.2
Cohabiting (married or partner)	74.2	69.9	56.4	60.7
Separated, divorced, widowed or other	17.2	19.1	24.1	18.2

**Table G.2 Respondents versus non-respondents, Wave 3 continued ...**

Wave 3	Productive %	Refusal %	Non-contact/ mover %	Other unproductive %
<i>Gross household income</i>				
Less than £10,400	12.9	21.5	26.5	17.5
£10,400 less than £16640	18.3	16.8	19.8	23.6
£16640 less than £26000	23.5	29.0	32.4	18.3
£26000 less than £36400	28.4	23.1	13.4	26.4
£36400 or more	16.8	9.7	7.9	14.2
<i>Government office region</i>				
North East	5.2	7.7	4.9	6.8
North West	15.4	15.0	13.2	9.0
Yorkshire and The Humber	10.7	9.0	10.9	3.2
East Midlands	8.5	6.9	9.4	7.5
West Midlands	11.3	13.8	15.3	12.8
East of England	10.8	6.7	9.3	15.6
London	12.0	17.8	15.4	18.1
South East	16.3	15.4	16.7	22.1
South West	9.8	7.6	5.0	5.0
<i>Deprivation index</i>				
1 Least deprived	25.8	21.1	13.0	20.2
2	21.2	18.7	8.6	14.7
3	22.6	24.7	23.1	22.2
4	16.6	19.1	28.2	20.6
5 Most deprived	13.8	16.4	27.1	22.4
<i>Base unweighted</i>	12382	568	293	202
<i>Base weighted</i>	12545	553	281	191

**Table G.3 Respondents versus non-respondents, Wave 4**

<b>Wave 4</b>	<b>Productive %</b>	<b>Refusal %</b>	<b>Non-contact/ mover %</b>	<b>Other unproductive %</b>
Owner occupied	76.1	67.0	50.5	60.1
Household does not have use of a motor vehicle	11.8	17.6	27.3	19.2
English is not the main/first language of the household	5.6	5.7	11.4	7.0
Lone parent household	22.1	28.7	39.3	38.8
MP has non-white ethnicity	10.9	11.2	18.2	12.7
Urban	79.5	75.8	85.6	80.5
Independent school	7.6	2.8	6.4	2.4
MP is male	14.4	15.8	16.3	14.2
Number of dependent children in the household (mean)	1.9	1.9	2.0	2.0
MPs age (mean)	44.4	43.1	42.0	42.6
<i>Main parents employment status</i>				
Employed	75.9	69.7	61.1	67.7
Unemployed	2.2	3.8	5.5	3.6
Economically inactive	21.9	26.6	33.4	28.6
<i>Family's NSSEC</i>				
Higher and lower managerial and professional occupations	41.7	34.1	29.0	30.2
Intermediate occupations	6.2	6.0	5.7	8.9
Small employers and own account workers	6.4	4.9	4.7	3.2
Lower supervisory and technical occupations	11.8	13.6	11.6	13.9
Semi-routine/routine occupations	18.9	22.1	22.5	19.3
Never worked/long term unemployed	15.0	19.3	26.6	24.5
<i>Main parents highest educational qualification</i>				
Degree or equivalent or higher education below degree	28.2	16.8	14.3	21.9
GCE A Level or equivalent GCSE grades A_C or equivalent	44.5	49.3	42.4	42.6
Qualifications at level 1 or below, or other qualifications	11.4	14.0	15.0	11.4
No qualifications	15.9	19.8	28.3	24.2
<i>Main parents marital status</i>				
Single (never married)	9.5	13.4	20.7	12.4
Cohabiting (married or partner)	70.2	61.6	50.7	53.8
Separated, divorced, widowed or other	20.3	25.0	28.6	33.8

**Table G.3 Respondents versus non-respondents, Wave 4 continued ...**

Wave 4	Productive %	Refusal %	Non-contact/ mover %	Other unproductive %
<i>Whether YP has a disability/long term illness or health problem</i>				
Has disability and schooling affected	6.4	6.4	5.6	12.4
Has disability, but schooling not affected	8.2	11.0	4.4	3.6
No disability or longstanding illness	85.4	82.6	90.0	84.0
<i>Gross household income</i>				
Less than £10,400	11.3	13.9	19.6	13.9
£10,400 less than £20800	21.6	31.2	38.6	31.9
£20800 less than £31200	20.5	20.8	14.6	21.4
£31200 less than £46800	21.6	19.7	18.9	20.5
£46800 or more	25.0	14.4	8.3	12.3
<i>Government office region</i>				
North East	5.1	6.2	7.2	3.3
North West	15.4	16.9	15.4	14.6
Yorkshire and The Humber	10.9	8.6	9.4	13.1
East Midlands	8.4	7.4	11.1	14.0
West Midlands	11.4	9.2	12.4	10.2
East of England	10.9	9.4	7.5	6.3
London	11.6	15.2	16.1	18.4
South East	16.4	18.9	14.9	9.5
South West	10.0	8.3	6.0	10.7
<i>Deprivation index</i>				
1 Least deprived	22.6	20.7	12.6	13.8
2	25.4	22.6	18.3	13.1
3	21.3	20.4	19.8	19.4
4	17.5	21.9	23.8	29.6
5 Most deprived	13.2	14.4	25.6	24.0
<i>Young Person's religion</i>				
None	42.4	49.2	47.5	44.9
Christian	47.4	40.3	39.8	40.7
Muslim	5.0	6.0	6.6	7.6
Other religion	5.2	4.4	6.1	6.8
<i>Base unweighted</i>	<i>11365</i>	<i>359</i>	<i>420</i>	<i>178</i>
<i>Base weighted</i>	<i>11602</i>	<i>358</i>	<i>363</i>	<i>157</i>

**Table G.4 Respondents versus non-respondents, Wave 5**

Wave 5	Productive %	Refusal %	Non-contact/ mover %	Other unproductive %
Owner occupied	77.7	69.7	50.2	67.6
Household does not have use of a motor vehicle	11.0	15.7	29.4	16.7
English is not the main/first language of the household	5.5	7.4	6.0	4.7
Lone parent household	24.3	29.8	44.3	32.4
MP has non-white ethnicity	12.4	12.5	15.2	8.7
Urban	79.0	83.0	87.5	79.9
Independent school	8.0	3.5	0.4	6.7
MP is male	49.4	65.6	52.9	56.4
Number of dependent children in the household (mean)	1.8	1.6	2.0	1.7
<i>Main parents employment status</i>				
Employed	77.8	72.1	63.7	72.0
Unemployed	1.9	2.5	4.5	2.7
Economically inactive	20.3	25.4	31.8	25.3
<i>Family's NSSEC</i>				
Higher and lower managerial and professional occupations	37.5	24.1	23.8	32.7
Intermediate occupations	8.7	7.6	4.4	7.0
Small employers and own account workers	9.8	10.6	10.4	10.5
Lower supervisory and technical occupations	7.4	15.4	10.4	9.6
Semi-routine/ routine occupations	19.5	20.4	20.1	17.6
Never worked/ long term unemployed	17.0	21.8	30.9	22.6
<i>Main parents highest educational qualification</i>				
Degree or equivalent or higher ed below degree	29.3	21.6	17.0	23.8
GCE A Level or equivalent GCSE grades A_C or equiv	45.0	45.7	40.6	40.8
Qualifications at level 1 or below, or other qualifications	10.8	14.0	16.1	15.1
No qualifications	14.9	18.6	26.3	20.2
<i>Main parents marital status</i>				
Single (never married)	8.2	9.3	19.1	10.2
Cohabiting (married or partner)	71.7	68.2	45.8	62.5
Separated, divorced, widowed or other	20.1	22.5	35.1	27.3

**Table G.4 Respondents versus non-respondents, Wave 5 continued ...**

Wave 5	Productive %	Refusal %	Non-contact/ mover %	Other unproductive %
<i>Gross household income</i>				
Less than £10,400	11.0	16.2	26.1	16.3
£10,400 less than £20800	20.0	24.0	27.5	24.3
£20800 less than £31200	19.6	23.6	20.4	17.4
£31200 less than £46800	21.8	16.5	16.4	20.2
£46800 or more	27.5	19.6	9.6	21.7
<i>Government office region</i>				
North East	4.9	7.3	6.1	7.2
North West	15.4	12.8	15.6	16.4
Yorkshire and The Humber	11.0	11.2	12.4	7.8
East Midlands	8.1	11.9	10.9	9.9
West Midlands	11.5	9.7	10.1	11.9
East of England	11.3	9.1	5.3	9.0
London	11.7	13.8	11.4	8.0
South East	16.3	17.6	13.5	19.1
South West	9.8	6.6	14.7	10.7
<i>Deprivation index</i>				
1 Least deprived	23.2	20.5	12.2	19.7
2	25.9	20.4	15.9	25.0
3	21.5	20.3	20.0	18.5
4	17.0	19.9	26.1	19.9
5 Most deprived	12.3	18.8	25.9	16.9
<i>Young Person's religion</i>				
None	36.6	41.0	44.5	44.7
Christian	54.7	47.4	47.1	48.9
Muslim	5.0	6.0	6.1	4.0
Other religion	3.7	5.6	2.2	2.4
<i>Whether Young Person has a disability/long term illness or health problem</i>				
Has disability and schooling affected	2.9	1.8	4.2	3.9
Has disability, but schooling not affected	5.1	5.0	4.2	4.1
No disability or longstanding illness	91.9	93.2	91.6	91.9
<i>Base unweighted</i>	<i>10036</i>	<i>364</i>	<i>440</i>	<i>413</i>
<i>Base weighted</i>	<i>10276</i>	<i>347</i>	<i>419</i>	<i>445</i>

**Table G.5 Respondents versus non-respondents, Wave 6**

Wave 6	Productive %	Refusal %	Non-contact/ mover %	Other unproductive %
Owner occupied	78.8	77.3	59.9	67.9
Household does not have use of a motor vehicle	10.4	12.4	23.7	13.9
English is not the main/first language of the household	5.5	4.0	3.9	5.9
Lone parent household	23.3	32.3	42.6	28.1
MP has non-white ethnicity	12.4	10.6	12.1	12.2
Urban	79.1	75.9	77.1	81.2
Independent school	8.2	4.8	6.4	5.3
MP is male	48.9	60.7	51.6	52.0
Number of dependent children in the household (mean)	1.8	1.6	1.8	1.9
<i>Main parents employment status</i>				
Employed	78.3	76.2	72.2	71.3
Unemployed	1.8	2.6	4.1	1.1
Economically inactive	19.9	21.2	23.7	27.6
<i>Family's NSSEC</i>				
Higher and lower managerial and professional occupations	38.5	28.7	26.5	28.7
Intermediate occupations	8.9	7.6	5.9	7.7
Small employers and own account workers	9.8	11.0	8.4	10.5
Lower supervisory and technical occupations	7.1	15.0	7.5	11.4
Semi-routine/ routine occupations	19.3	19.6	26.2	19.0
Never worked/ long term unemployed	16.4	18.0	25.5	22.6
<i>Main parents highest educational qualification</i>				
Degree or equivalent or higher ed below degree	30.2	20.6	20.2	20.4
GCE A Level or equivalent GCSE grades A_C or equivalent	45.1	47.6	40.2	44.9
Qualifications at level 1 or below, or other qualifications	10.5	12.6	12.3	17.0
No qualifications	14.2	19.2	27.3	17.6
<i>Main parents marital status</i>				
Single (never married)	8.0	7.5	12.4	9.0
Cohabiting (married or partner)	72.9	66.9	50.1	66.9
Separated, divorced, widowed or other	19.1	25.7	37.5	24.1

**Table G.5 Respondents versus non-respondents, Wave 6 continued ...**

Wave 6	Productive %	Refusal %	Non-contact/ mover %	Other unproductive %
<i>Gross household income</i>				
Less than £10,400	10.5	16.9	15.8	17.1
£10,400 less than £20800	19.8	15.2	30.4	21.5
£20800 less than £31200	19.7	18.2	17.7	22.4
£31200 less than £46800	21.9	24.8	18.9	20.0
£46800 or more	28.2	24.9	17.1	19.0
<i>Government office region</i>				
North East	4.7	7.1	7.4	5.9
North West	15.2	17.9	19.5	17.2
Yorkshire and The Humber	10.8	11.2	11.3	15.8
East Midlands	8.2	7.9	7.6	7.5
West Midlands	11.5	9.1	11.0	14.0
East of England	11.4	11.0	11.2	8.1
London	11.9	10.6	9.8	8.6
South East	16.6	15.0	11.3	15.2
South West	9.9	10.1	10.8	7.7
<i>Deprivation index</i>				
1 Least deprived	23.9	22.1	12.4	15.8
2	26.0	22.4	25.4	27.7
3	21.6	22.3	22.1	16.8
4	16.7	20.1	21.0	22.1
5 Most deprived	11.9	13.1	19.0	17.7
<i>Young Person's religion</i>				
None	36.1	37.6	45.1	39.8
Christian	55.1	54.4	49.4	49.6
Muslim	5.0	5.6	4.1	6.6
Other religion	3.8	2.4	1.3	4.0
<i>Whether Young Person has a disability/long term illness or health problem</i>				
Has disability and schooling affected	2.9	3.3	1.6	3.4
Has disability, but schooling not affected	5.2	4.4	4.7	4.5
No disability or longstanding illness	91.9	92.4	93.8	92.1
<i>Base unweighted</i>	<i>9119</i>	<i>295</i>	<i>342</i>	<i>265</i>
<i>Base weighted</i>	<i>9331</i>	<i>297</i>	<i>354</i>	<i>278</i>

**Table G.6a**      **Logistic regression results**

Dependent variable: 0=responded, 1=refused	Wave 2	Wave 3	Wave 4
<i>Employment status of main parent</i> <i>p=0.009</i>			
Employed	1		
Unemployed	1.42 (0.8,2.53)		
Economically inactive	1.35 (1.11,1.65)		
<i>Highest educational qualification of main parent</i> <i>p=0.012</i> <i>p=0.024</i> <i>p=0.024</i>			
Degree or equivalent or higher education below degree	1	1	1
GCE A Level or GCSE grades A_C or equivalent	1.28 (1.01,1.62)	1.35 (1.02,1.8)	1.63 (1.19,2.24)
Qualifications at level 1 or below, or other qualifications	1.54 (1.14,2.09)	1.5 (1.04,2.17)	1.67 (1.07,2.6)
No qualifications	1.54 (1.17,2.03)	1.64 (1.19,2.27)	1.59 (1.06,2.39)
<i>Main parents sex</i> <i>p=0.013</i>			
Male	1		
Female	0.76 (0.61,0.94)		
<i>Gross household income bands</i> <i>p&lt;0.001</i> <i>p&lt;0.001</i> <i>p&lt;0.001</i>			
Less than £10,400	1	1	1
£10,400 less than £16640 (to less than £20,800 for Waves 3 and 4)	0.76 (0.56,1.04)	0.57 (0.36,0.89)	1.18 (0.77,1.81)
£16640 less than £26000 (£20,800 to less than £33,800 for Wave 3, and £20,800 to less than £31,200 for Wave 4)	0.86 (0.63,1.17)	0.79 (0.53,1.18)	0.84 (0.52,1.33)
£26000 less than £36400 (£33,800 to less than £55,000 for Wave 3 and £31200 to less than £46800 for Wave 4)	0.67 (0.47,0.96)	0.54 (0.36,0.8)	0.79 (0.48,1.3)
£36400 or more (£55,000 or more for Wave 3 and £46,800 or more for Wave 4)	0.86 (0.61,1.19)	0.4 (0.23,0.7)	0.54 (0.31,0.94)
Missing	1.39 (1.07,1.8)	1.07 (0.78,1.47)	1.65 (1.05,2.59)
<i>Government Office Region</i> <i>p=0.002</i>			
London		1	
North East		0.7 (0.46,1.07)	
North West		0.58 (0.35,0.95)	
Yorkshire and The Humber		0.56 (0.36,0.87)	
East Midlands		0.93 (0.6,1.44)	
West Midlands		0.48 (0.28,0.83)	
East of England		1.03 (0.7,1.53)	
South East		0.78 (0.51,1.17)	
South West		0.59 (0.37,0.95)	
<i>Main parents age</i> <i>p=0.014</i>			
		0.98 (0.96,1)	
<i>Urban/rural index</i> <i>p=0.033</i>			
Rural			1
Urban			0.73 (0.55,0.97)

**Table G.6b Logisitic regression results**

Dependent variable: 0=responded, 1=refused	Wave 5	Wave 6
<i>Young persons sex</i>	$p<0.001$	$p<0.001$
Male	1	1
Female	0.47 (0.36,0.62)	0.58 (0.44,0.78)
<i>Households NSSEC classification</i>	$p=0.002$	$p=0.002$
Higher and lower managerial and professional occupations	1	1
Intermediate occupations	1.21 (0.71,2.07)	0.97 (0.56,1.7)
Small employers and own account workers	1.54 (0.94,2.52)	1.5 (0.91,2.49)
Lower supervisory and technical occupations	2.65 (1.68,4.19)	2.65 (1.7,4.13)
Semi-routine/routine occupations	1.22 (0.81,1.85)	1.22 (0.81,1.85)
Not currently working	1.59 (1.08,2.36)	1.41 (0.93,2.14)
<i>Number of dependent children in the household</i>	$p=0.004$	$p=0.039$
	0.83 (0.73,0.94)	0.86 (0.74,0.99)
<i>Quintiles of the Index of multiple deprivation (2004)</i>	$p=0.012$	
1 Least deprived	1	
2	0.85 (0.54,1.32)	
3	1.21 (0.8,1.82)	
4	1.42 (0.94,2.13)	
5 Most deprived	1.92 (1.2,3.08)	
<i>Government Office Region</i>	$p=0.007$	
London	1	
North East	1.25 (0.67,2.33)	
North West	0.65 (0.38,1.12)	
Yorkshire and The Humber	0.8 (0.44,1.44)	
East Midlands	1.54 (0.92,2.57)	
West Midlands	0.7 (0.4,1.23)	
East of England	0.81 (0.45,1.48)	
South East	1.29 (0.8,2.08)	
South West	0.62 (0.34,1.13)	
<i>Young persons self completion questionnaire</i>	$p=0.001$	
Respondent accepted self completion	1	
Respondent refused self completion or interviewer administered the self completion	2.66 (1.46,4.82)	
<i>Whether lone parent family or not</i>		$p=0.006$
No		1
Yes		1.52 (1.13,2.04)

**Table G.6c**      **Logistic regression results**

Dependent variable: 0=responded, 1=non-contact/mover			
	Wave 2	Wave 3	Wave 4
<i>Accommodation is owner-occupied</i>			
	<i>p</i> <0.001	<i>p</i> <0.001	<i>p</i> <0.001
No	1	1	1
Yes	0.54 (0.44,0.65)	0.37 (0.27,0.51)	0.51 (0.37,0.7)
<i>Fieldwork agency</i>			
	<i>p</i> <0.001		
BMRB	1		
NOP	0.43 (0.35,0.54)		
MORI	0.66 (0.5,0.88)		
<i>Whether lone parent family or not</i>			
	<i>p</i> <0.001		
No	1		
Yes	1.73 (1.45,2.08)		
<i>Government Office Region</i>			
	<i>p</i> =0.026		
London	1		
North East			
	0.68 (0.41,1.14)		
North West			
	0.83 (0.6,1.15)		
Yorkshire and The Humber			
	0.7 (0.49,0.98)		
East Midlands			
	0.65 (0.43,0.98)		
West Midlands			
	0.54 (0.38,0.77)		
East of England			
	1.12 (0.68,1.85)		
South East			
	0.77 (0.54,1.08)		
South West			
	0.64 (0.4,1.02)		
<i>Quintiles of the Index of multiple deprivation (2004)</i>			
	<i>p</i> =0.008	<i>p</i> =0.028	
1 Least deprived	1	1	
2	0.91 (0.7,1.19)	0.61 (0.34,1.1)	
3	1.13 (0.84,1.51)	1.27 (0.8,2.01)	
4	1.47 (1.09,1.97)	1.48 (0.92,2.38)	
5 Most deprived	1.18 (0.86,1.63)	1.48 (0.89,2.48)	
<i>Main parents sex</i>			
	<i>p</i> =0.004		
Male	1		
Female	0.73 (0.59,0.91)		
<i>Main parents age</i>			
	<i>p</i> =0.003	<i>p</i> =0.017	<i>p</i> <0.001
	0.98 (0.96,0.99)	0.97 (0.94,0.99)	0.96 (0.93,0.98)

**Table G.6c** Logistic regression results continued ...

Dependent variable: 0=responded, 1=non-contact/mover	Wave 2	Wave 3	Wave 4
<i>Independent school (Wave 1)</i>	<i>p=0.011</i>		<i>p=0.042</i>
No	1		1
Yes	1.73 (1.13,2.63)		1.99 (1.03,3.86)
 <i>Ethnicity of main parent</i>	 <i>p&lt;0.001</i>		 <i>p=0.004</i>
Non-white	1		1
White	0.59 (0.47,0.73)		0.68 (0.52,0.88)
 <i>Urban/rural index</i>		 <i>p=0.050</i>	
Rural		1	
Urban		1.72 (1,2.96)	
 <i>Gross household income bands</i>		 <i>p=0.024</i>	 <i>p&lt;0.001</i>
Less than £10,400		1	1
£10,400 less than £16640 (to less than £20,800 for Waves 3 and 4)		0.66 (0.37,1.2)	1.17 (0.8,1.73)
£16640 less than £26000 (£20,800 to less than £33,800 for Wave 3, and £20,800 to less than £31,200 for Wave 4)	1.02 (0.58,1.79)	0.63 (0.38,1.05)	
£26000 less than £36400 (£33,800 to less than £55,000 for Wave 3 and £31200 to less than £46800 for Wave 4)	0.45 (0.23,0.89)	0.9 (0.55,1.47)	
£36400 or more (£55,000 or more for Wave 3 and £46,800 or more for Wave 4)	0.59 (0.25,1.39)	0.37 (0.19,0.69)	
Missing		1.12 (0.71,1.77)	1.12 (0.73,1.73)
 <i>Whether household has use of a motor vehicle</i>			 <i>p=0.049</i>
No			1
Yes			0.72 (0.52,1)

**Table G.6d**      **Logistic regression results**

Dependent variable: 0=responded, 1=non-contact/mover		
	Wave 5	Wave 6
<i>Accommodation is owner-occupied</i>		
	<i>p</i> <0.001	<i>p</i> =0.006
No	1	1
Yes	0.49 (0.37,0.64)	0.66 (0.5,0.89)
<i>Whether lone parent family or not</i>		
	<i>p</i> <0.001	<i>p</i> <0.001
No	1	1
Yes	1.63 (1.27,2.08)	2.01 (1.49,2.71)
<i>Urban/rural index</i>		
	<i>p</i> =0.034	
Rural	1	
Urban	1.51 (1.03,2.21)	
<i>Quintiles of the Index of multiple deprivation (2004)</i>		
	<i>p</i> =0.031	
1 Least deprived	1	
2	0.97 (0.61,1.53)	
3	1.24 (0.79,1.95)	
4	1.55 (1.01,2.4)	
5 Most deprived	1.79 (1.13,2.82)	
<i>Government Office Region</i>		
	<i>p</i> =0.016	
London	1	
North East	1.38 (0.78,2.42)	
North West	1.15 (0.73,1.79)	
Yorkshire and The Humber	1.29 (0.8,2.08)	
East Midlands	1.93 (1.14,3.29)	
West Midlands	1.1 (0.67,1.8)	
East of England	0.73 (0.43,1.26)	
South East	1.49 (0.93,2.4)	
South West	2.07 (1.17,3.68)	
<i>Gross household income bands</i>		
	<i>p</i> <0.001	
Less than £10,400	1	
£10,400 less than £16640 (to less than £20,800 for Waves 3 and 4)	0.77 (0.54,1.1)	
£16640 less than £26000 (£20,800 to less than £33,800 for Wave 3, and £20,800 to less than £31,200 for Wave 4)	0.85 (0.58,1.26)	
£26000 less than £36400 (£33,800 to less than £55,000 for Wave 3 and £31200 to less than £46800 for Wave 4)	0.8 (0.51,1.25)	
£36400 or more (£55,000 or more for Wave 3 and £46,800 or more for Wave 4)	0.45 (0.26,0.78)	
Missing	1.44 (1,2.07)	
<i>Fieldwork agency</i>		
	<i>p</i> =0.036	
BMRB	1	
NOP	0.85 (0.58,1.26)	
MORI	0.8 (0.51,1.25)	

**Table G.6d** Logistic regression results continued

Dependent variable:

0=responded, 1=non-contact/mover

Wave 5

Wave 6

	Wave 5	Wave 6
<i>Highest educational qualification of main parent</i>		<i>p=0.001</i>
Degree or equivalent or higher education below degree		1
GCE A Level or equiv GCSE grades A_C or equiv	1.21 (0.86,1.71)	
Qualifications at level 1 or below, or other qualifications	1.45 (0.91,2.29)	
No qualifications	2.28 (1.46,3.55)	
<i>Young Person's religion</i>		<i>p=0.006</i>
None		1
Christian	0.75 (0.59,0.97)	
Muslim	0.5 (0.28,0.89)	
Other religion	0.32 (0.13,0.81)	

## How and why response rates changed over time

For Waves 2 to 4, these categories (refusals, non-contacts etc) refer to the main parent, as they had to provide consent for the young person to respond. Thus for these Waves there is also an extra category of ‘young person incomplete interview’ where the main parent responded but the young person did not (there is no extra information on why the young person did not respond). In Waves 5 and 6 this category is null as the other categories (refusals, non-contacts etc) refer to the young person directly (as they were approached directly without requiring consent from the main parent).

**Table G.7 Non-responders (those who responded at the previous Wave)**

	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
	%	%	%	%	%
Refusal/opt-out	37	47	32	29	32
No contact	20	11	8	14	15
Broken appointment	11	12	9	5	7
Mover/address inaccessible or unable to locate	23	15	30	23	24
Other unproductive	1	5	7	29	22
Young person incomplete interview (but another module completed)	8	10	15	0	0
Total number of unproductives	2243	1114	1084	1150	941

**Table G.8 Response rates by ethnicity for each Wave**

Wave	White %	Mixed %	Asian %	Black %	Chinese/Other %
2	87.2	82.9	84.3	76.5	77.0
3	92.5	88.6	91.6	85.3	88.1
4	92.3	89.4	91.1	83.6	88.1
5	89.9	87.6	90.6	87.9	87.0
6	90.6	89.4	92.4	88.5	96.0
Responded in all Waves	61.1	53.0	59.4	42.7	55.2

**Table G.9 Response rates by Government Office Region for each Wave**

Wave	North East	North West	Yorkshire and the Humber	East Midlands	West Midlands	East of England	London	South East	South West
	%	%	%	%	%	%	%	%	%
2	88.0	84.8	87.7	86.5	86.4	85.5	80.0	87.7	88.8
3	90.8	93.3	93.2	92.7	89.8	93.4	88.2	91.3	94.5
4	91.2	91.0	92.2	90.4	91.9	93.1	87.8	92.2	92.7
5	86.4	89.4	90.0	87.2	90.2	92.8	90.3	90.1	89.0
6	87.3	88.9	89.6	91.3	91.5	91.6	91.8	92.0	90.6
All waves	55.9	58.2	61.7	58.1	59.3	63.5	51.8	61.6	63.5

**Table G.10 Predictors of movers or non-contacts at Wave 2**

	Movers	Non-contacts
<i>Accommodation is owner-occupied</i>		
No	1	
Yes	0.35 (0.27,0.45)	
<i>Ethnicity of main parent</i>		
Non-white	1	1
White	0.51 (0.40,0.66)	0.48 (0.35,0.65)
<i>Main parents sex</i>		
Male	1	
Female	0.69 (0.52,0.92)	
<i>Whether lone parent family or not</i>		
No	1	1
Yes	1.66 (1.32,2.08)	1.85 (1.39,2.46)
<i>Main parents age</i>		
	0.96 (0.95,0.98)	
<i>Independent school (Wave 1)</i>		
No		1
Yes		1.96 (1.11,3.45)
<i>Government Office Region</i>		
London		1
North East		0.65 (0.31,1.35)
North West		0.71 (0.44,1.13)
Yorkshire and The Humber		0.45 (0.26,0.79)
East Midlands		0.48 (0.27,0.85)
West Midlands		0.38 (0.22,0.65)
East of England		1.04 (0.46,2.36)
South East		0.45 (0.26,0.80)
South West		0.41 (0.21,0.81)

## Appendix H      Details of fieldwork years for other studies

**Olympic Babies**, which is part of the Birth Cohort Facility project, is a gateway to exploring how our careers, health, wellbeing and personal relationships are forged from childhood through to retirement. It has been funded by the Medical Research Council, and will be designed and launched as a major new Birth Cohort Study in 2012. The fieldwork is planned to commence in spring 2012. The form and mode of the interview is still under discussion.

The fifth sweep of the **Millennium Cohort Study** (MCS) is scheduled to take place in 2012 when the cohort children will be aged 11 and completing their final year of primary school. It is provisionally assumed that the fieldwork will take place from January to December 2012 (as in previous sweeps) in the UK, with fieldwork in England taking place from January to August. MCS 5 will be carried out in the cohort child's home and will comprise of a paper self completion questionnaire for the cohort child and face to face interviews with their mother and her resident partner. In addition, the cohort child's teacher will be asked to complete a self completion questionnaire which is sent to the child's school.

The next **BCS70** sweep will take place in 2012 when the cohort members reach the age of 42, and will be a face to face interview with cohort members at their homes. Details of fieldwork timings within the year are not confirmed but are likely to be May to November.

**Understanding Society** will have two Waves in field in 2012. Wave 3 will be in its second year and Wave 4 in its first year (each Wave takes 24 months to complete). Samples for both Waves will be issued on a monthly basis from January to December. In each eligible household, every adult (age 16 and over) will be interviewed face to face and asked to complete a self completion questionnaire. In addition there will be a separate self completion questionnaire for all children aged 10 to 15 living in the household. Interviews will take place in cohort members' homes.

# Appendix I Expert Panel review of content of LSYPE

An Expert Panel was convened to review the first LSYPE cohort questionnaires, and identify those questions most likely to be susceptible to mode effects according to current understanding of measurement differences across modes.

Members of the Expert Panel were:

- (a) Aleks Collingwood, Senior Researcher (Analyst)
- (b) Michelle Gray, Survey Methodologist (Questionnaire Development and Testing)
- (c) Steven Hope, Survey Methodologist (Data Collection Methodology)
- (d) Gerry Nicolaas, Deputy Director of the Survey Methods Unit
- (e) Andy Ross, Senior Researcher (Analyst)

Aleks Collingwood and Andy Ross have considerable experience of analysing data from the first LSYPE cohort and are therefore familiar with the data and how it is used. Gerry Nicolaas, Steven Hope and Michelle Gray are survey methodologists with experience of designing questions for mixed mode studies.

The Expert Panel identified a range of question types and formats that have been used in the first LSYPE cohort and reviewed how these questions would perform across different modes; i.e. CAWI, CATI, CAPI and CASI. The conclusions from this review are summarised in the following table along with some suggestions on how to minimise mode effects.

Table I.1 Expert Panel review of content						
Question types & formats in LSYPE	LSYPE Examples	Comments	CAWI	CATI	CAPI (CASI)	Suggestion
Simple factual, non-sensitive questions with short response lists	Do you have a mobile phone of your own? 1. Yes 2. No	On the whole, no mode effects if question wording is identical in all modes.	✓	✓	✓	See Dillman's unimode principles (Dillman et al, 2009)
Sensitive questions	Have you ever smashed, slashed or damaged public property or something in a public place? 1. Yes 2. No 3. Don't know 4. Don't want to answer	There is evidence that sensitive questions are prone to social desirability effects in interviewer modes, although this has been found to vary by type of question and degree of sensitivity.	✓		✓ (CASI)	Option for the most sensitive questions is to collect responses using self completion modes (i.e. CAWI & CASI). Explore feasibility of sending self completion questionnaire by post/email to telephone respondents or using touchtone data entry (TDE). However, this approach may be time consuming and costly to implement.
Open questions	Why do you want to leave full time education?	Interviewers can encourage respondents to provide full answers using probing techniques, something that is not possible in self-completion modes. Nevertheless, the evidence regarding the quality of data collected using open questions in different modes is mixed.		✓	✓	Limit the use of open questions

Table I.1 Expert Panel review of content continued ...

Question types & formats in LSYPE	LSYPE Examples	Comments	CAWI	CATI	CAPI (CASI)	Suggestion
Show cards in CAPI	Looking at this card, in what way(s) do you use your computer at home for school work? 1. Word processing 2. Spreadsheets 3. Graphics 4. Other sorts of packages on the computer 5. E-mails 6. Use the web to find information on homework topics 7. Something else	On the whole, comparable responses can be achieved in CAPI using show cards and CAWI (both use visual presentation of response list). Visual presentation of response list is not possible in CATI which may result in more recency effects compared to more primacy effects in CAPI with show cards and CAWI.	✓		✓ (show card)	Minimise use of show cards by reducing the list of response options for all modes. When show cards are being used in CAPI, consider randomising the order of the response list in all modes.
Agree-Disagree Scales  Non-sensitive agree-disagree questions without a middle category	There's a good range of shops in my local area 1. Strongly agree 2. Agree 3. Disagree 4. Strongly Disagree	Problematic in all modes because it is a cognitively complex task and may be susceptible to acquiescence bias. There is some evidence to suggest that there will be a primacy effect when a scale is presented visually and a recency effect when presented aurally.  Primacy effect.....  Recency effect.....	✓		✓ (CASI)  ✓	Avoid use of agree-disagree scales and use alternative formats; e.g. questions with Item Specific (IS) response options. To overcome primacy and recency effects, it is possible to randomise the order of the scale. However, this may not work for agree-disagree scales because ordering the response options from strongly disagree to strongly agree is atypical and may confuse respondents resulting in measurement error.

Table I.1 Expert Panel review of content continued ...

Question types & formats in LSYPE	LSYPE Examples	Comments	CAWI	CATI	CAPI (CASI)	Suggestion
Non-sensitive agree-disagree questions with a middle category	No first LSYPE cohort question of this type	There is some evidence to suggest that CAWI respondents are more likely to select middle categories than CAPI and CATI respondents.		✓	✓	The usual recommendation is to include a middle category unless there are persuasive reasons to omit it. Explore the impact of adding an explicit 'don't know' category across all modes when using middle categories.
Sensitive agree-disagree questions	The work I do in lessons is a waste of time. 1. Strongly agree 2. Agree 3. Disagree 4. Strongly disagree	CAPI and CATI respondents are more likely than CAWI respondents to select socially desirable response.	✓		✓ (CASI)	Only use in self completion modes (i.e. CAWI & CASI). Explore feasibility of sending self completion questionnaire by post/email to telephone respondents or using touchtone data entry (TDE).
Batteries of questions using the same response list	For the next few questions, you will be given some reasons why people might choose to study a particular subject at university. For each of these please say how important it was for you when you were thinking about what subject you wanted to do at university. Because you need a degree in this subject to get a specific job or career you want to do after university. Was this.... Very important Fairly important Not very important Not at all important	These questions are more susceptible to non-differentiation in self completion modes.		✓	✓	Avoid long sequences of questions using the same rating scale.

Table I.1 Expert Panel review of content continued ...

Question types & formats in LSYPE	LSYPE Examples	Comments	CAWI	CATI	CAPI (CASI)	Suggestion
Open questions with long list of response options for interviewer coding	<p>What is your favourite subject at school?</p> <p>PROMPT TO PRECODE AND CODE ONE ONLY</p> <ol style="list-style-type: none"> <li>1 Mathematics</li> <li>2 Science (biology, chemistry, physics)</li> <li>3 Design and Technology</li> <li>4 ICT/Information and Communication Technology/Computing</li> <li>5 Home Economics</li> <li>6 History</li> <li>7 Geography</li> <li>8 Physical Education (p.e.)/Games/Sport (including individual sports)</li> <li>9 Business studies or economics</li> <li>10 Humanities, social or vocational studies</li> <li>11 Art</li> <li>12 English</li> <li>13 Modern languages (e.g. French, German, Spanish)</li> <li>14 Music</li> <li>15 Drama or media/film/television studies or communication studies</li> <li>16 Religious studies</li> <li>17 Sex and health education</li> <li>18 Citizenship</li> <li>19 Other (specify)</li> </ol>	<p>Presentation of response options in CAWI may produce different responses than the open question format in CAPI &amp; CATI with interviewer coding.</p>		✓	✓	<p>Consider instructing interviewers not to prompt but to accept respondent's first answer.</p> <p>Equivalent CAWI format would be an open question with office-coding of responses.</p>
Multi-coded questions	<p>Do you ever use your mobile phone for any of these things?</p> <p>READ OUT AND CODE ALL THAT APPLY</p> <ol style="list-style-type: none"> <li>1. Accessing web sites</li> <li>2. Emailing someone</li> <li>3. Sending picture messages to someone</li> <li>4. Playing games</li> <li>5. None of these</li> </ol>	<p>Respondents tend to record fewer options in self completion modes than in interviewer modes.</p>		✓	✓	<p>Consider using forced choice questions for each response option in all modes. Note that this can add to the length of the interview and can be quite tedious for long lists of items in CAPI and CATI.</p>

Table I.1 Expert Panel review of content continued ...

Question types & formats in LSYPE	LSYPE Examples	Comments	CAWI	CATI	CAPI (CASI)	Suggestion
Ordinal scales (e.g. frequencies)	<p>In the last 12 months would you say your health has been very good, fairly good, not very good or not good at all?</p> <ol style="list-style-type: none"> <li>1. Very good</li> <li>2. Fairly good</li> <li>3. Not very good</li> <li>4. Not good at all</li> </ol> <p>(Example. Could be changed to a sequence of branching questions:            In the last 12 months would you say your health has been good or not good?            IF GOOD: Would you say your health has been fairly good or very good?            IF NOT GOOD: Would you say your health has been not very good or not good at all?)</p>	<p>In telephone surveys, ordinal scales are often changed into a sequence of two or more branching questions in order to reduce the cognitive burden for the telephone respondent. However, there is evidence to suggest that this will change how respondents answer questions compared to offering all categories at once.</p>	✓		✓ (show card)	<p>If the use of a single ordinal scale is too burdensome in a telephone interview, consider using a sequence of two or more branching questions across all modes. There is some evidence to suggest that this can improve measurement reliability and validity (for example, see Malhotra, 2009). Dillman recommends not using a sequence of branching questions in mixed mode surveys but to reduce the number of categories offered so that branching is not required.</p>
Use of middle categories	<p>What about people like yourself? How fairly do you think people like yourself are treated by government, either nationally or your local council, in Britain today? Would you say people like you are usually treated...</p> <ol style="list-style-type: none"> <li>1. Very fairly</li> <li>2. Quite fairly</li> <li>3. Neither fairly or unfairly</li> <li>4. A little unfairly</li> <li>5. Very unfairly?</li> </ol>	<p>There is a greater tendency for middle categories to be selected in self completion modes than in interviewer modes.</p>		✓	✓	<p>The usual recommendation is to include a middle category unless there are persuasive reasons to omit it. Explore the impact of adding an explicit 'don't know' category across all modes when using middle categories.</p>

Table I.1 Expert Panel review of content continued ...

Question types & formats in LSYPE	LSYPE Examples	Comments	CAWI	CATI	CAPI (CASI)	Suggestion
Spontaneous 'don't know' in CAPI & CATI	<p>What, if any, is your religion?</p> <ol style="list-style-type: none"> <li>1. None</li> <li>2. Christian</li> <li>3. Buddhist</li> <li>4. Hindu</li> <li>5. Jewish</li> <li>6. Muslim</li> <li>7. Sikh</li> <li>8. Another religion</li> </ol> <p>(Spontaneous only: Don't know ) Refused)</p>	<p>It is common practice in telephone and face to face interviews not to provide respondents with an explicit 'don't know' option but the interviewer can record this answer if the respondent spontaneously says 'don't know'. In self completion modes, the 'don't know' option is either an explicit response option or it is omitted altogether. Treating 'don't know' differently in different modes can result in different rates of 'don't know' across the modes.</p>		✓	✓	<p>Spontaneous 'don't know' can only be used in CATI and CAPI. Consider use of explicit 'don't know' across all modes. For some questions this may increase survey satisficing, particularly in CAWI and CASI. For those questions where 'don't know' could be a valid response (e.g. opinion questions), consider using filter questions to determine interest and knowledge in a topic and then drop the 'don't know' option from the main question.</p>
Use of interviewer instructions, clarifications, etc	<p>Can I check, is English your first or main language? INTERVIEWER: If 'Yes', Probe - 'Is English the only language you speak or do you speak any other languages, apart from languages you may be learning at school as part of your studies?'</p> <ol style="list-style-type: none"> <li>1. Yes - English only,</li> <li>2. Yes - English first/main and speaks other languages,</li> <li>3. No, another language is respondent's first or main language,</li> <li>4. Respondent is bilingual</li> </ol>	<p>It is common practice to provide interviewers with additional information that can be used if necessary to improve the quality of information from the respondents. This can result in differences across modes in a study that uses self completion modes alongside interviewer modes.</p>		✓	✓	<p>Where possible, all instructions and clarifications should be added to the question for all modes or excluded from all modes.</p>

## Appendix J Theories of exchange

Chapter Six explores the use of incentives. Theories that can be applied to interpreting the effects of incentives, which were reviewed as part of this development work, are discussed below.

Theories that can be applied to interpreting the effects of incentives

Social exchange theory provides the most commonly used theoretical framework for interpreting the effects of incentives on survey participation (Dillman 1978; Groves, Cialdini & Couper 1992). When faced with a survey request, individuals will provide help in proportion to the gain that is expected or has been received (in this case an incentive) in order to maintain equity in the relationship. Related to this is also the norm of reciprocity (Gouldner 1960) which has been used to explain why receiving an incentive leads to feeling of obligation to cooperate with a survey request next time. However some regard the relationship between incentives and survey cooperation as more complex than this and argue that intrinsic motivation (e.g. altruism, civic duty or interest) plays an important role in overall performance on task (Deci et al, 1999).

Are these theories relevant for children?

Much of the literature suggests that behavioural theories of the kind above can be applied to children. Sutter & Kocher (2003) identify that although trust increases significantly with age, reciprocal behaviour can be seen as early as age 8. This is supported by Harbaugh & Krause (2000) who find no difference in altruistic preferences between 6 to 12 year olds and adults in a public good game. Further work by Harbaugh & Krause (2000) has adapted economic experiments to study children's interactive decision making using real monetary incentives. They argue that children's behaviour is very close to that of adults if faced with economic choices and they behave as economic theory would predict by rationalising to come up with a decision. Interestingly this is true even if they are too young to calculate probabilities or expected returns.

Pierce & Cameron (2002) have applied research on rewards and intrinsic motivation to an educational setting by highlighting that teachers use praise, gold stars, and other such incentives to promote learning. They also stress that some parents offer children rewards for doing well at school, or for sports accomplishments. In this context, reward for performance is seen as a part of everyday life, and they argue against the notion that rewards undermine performance and motivation. Pierce & Cameron (2002) also comment on the use of rewards over time. They summarise five operant studies where respondents were observed over a number of sessions (baseline) in a non-reward phase. A reward was then presented over a number of sessions, before the reward was taken away again and performance assessed in the absence of the reward. Intrinsic motivation was measured as the difference in performance between the pre and post reward phases. They found that there was no detrimental effect of reward on performance. Performance in the post-reward phase recovered to a level that matched or exceeded the pre-reward phase. They also

found no decline in performance following removal of a reward recovers – meaning that any detrimental effects are temporary.

This work shows it is possible to introduce a reward at any given stage of a task and then remove it without significantly affecting overall performance thereafter. This could relate, in theory, to introducing an incentive at a current Wave of interviewing on a longitudinal survey and not offering it at subsequent Waves.

- An article by Collins et al (2008) supports the use of incentives with children in order to increase participation in out-of-school time programs. They propose four tips regarding the use of incentives for out-of-school time programs. In theory, these could also be generalised to the request for survey participation.
  - Ask program respondents for ideas about incentives;
  - Introduce incentives immediately after goals have been reached so they draw a correlation between the desired behaviour and reward;
  - Gain community support by asking local businesses to fund or provide incentives, or offer promotions;
  - Use incentives sparingly. Some research suggests that if incentives are used too often, program respondents start to rely on them alone as a motivation for attending (Benabou & Tirole 2003). Singer et al (1999) also found that students who cooperated with a survey request after receiving a small gift perceived themselves as motivated by interest, while those given \$10 attributed their participation to the incentive.

## Appendix K Monetary Incentives

Examples of studies with a monetary incentive

### *British Household Panel Study (BHPS)*

The British Household Panel Study comprises an annual interview which started in 1991. For the first five years, longitudinal members were offered £5 for an adult interview. In 1996 this rose to £7 and then to £10 from 2005. Young people aged 11-15 are incentivised to fill in a self completion questionnaire. Payment for the youth self completion rose from £3 in 1994 to £5 from 2005.

For adults and young people, the incentive is sent in advance if they have been interviewed at the previous Wave, otherwise interviewers have spare vouchers to give at the end of the interview for prior non-respondents or any new household members.

### *English Longitudinal Study of Ageing (ELSA)*

ELSA started in 2002, with interviews every two years. At each Wave, longitudinal members have been offered £10 on completion of an interview. Restrictions imposed by the Medical Research and Ethics Committee (MREC) prevented the study from switching to the use of unconditional incentives. However, the study adopts the use of differential incentives for non-respondents at a given Wave, with £20 offered to those who refuse the initial contact.

International studies

### *Longitudinal Study of Income Dynamics (PSID)*

PSID originally offered \$5 when it started in 1968 and this rose to \$20 by 1995. In 1999 the study changed from an annual to a biennial study with a longer interview when the incentive was increased to \$40, rising again to \$60 in 2005. Current practice is for PSID to pay an incentive of around \$1 per interview minute, with a cheque being sent out roughly one week after the interview.

### *National Longitudinal Surveys of Youth (NLSY)*

NLSY has seen an increase in incentive payment from \$10 when it first started in 1979 to \$20 by 1996. In 1998 some households were offered a bonus of between \$100 and \$150 in reaction to response rate concerns. However, the subsequent round of interviewing was reported to suffer when these payments were not replicated. In 2002 the base rate for interview rose to \$40, but with increased payments being offered to those who agreed to phone in to do an interview over the telephone.

### *Health and Retirement Study (HRS)*

The HRS conducts interviews every two years, and although these are mainly telephone interviews, some Waves have used face to face interviewing. Between 1992 and 2002 longitudinal members were sent an unconditional payment of \$20 to participate in an interview, increasing to \$40 in 2004. In 2006, a sub-sample was asked to do an expanded face to face interview so the incentive rate was increased to \$50 for this group. The HRS is also well known for the 'end-game' strategy adopted at the first Wave when large financial bonuses of \$100 were offered to reluctant respondents for an immediate yes/no decision to participate.

### *Household Income and Labour Dynamics in Australia (HILDA) survey*

The HILDA survey adopted a slightly different approach to its first four Waves by providing a financial incentive to the household rather than the individual. If all household members were interviewed a full productive rate of Aus\$50 was administered following the interview. In contrast, those households with missing individual interviews were given a partial fee of Aus\$20. From 2005, there was a switch to incentivising the individual (rather than the household) with HILDA respondents receiving \$25 per individual with a bonus of \$25 to the household reference person. On average, it takes about 6 weeks following the interview for respondents to receive their cheque.

### Examples of studies that have increased the incentive after Wave 1

#### *Health and Retirement Study*

An incentive experiment was conducted during the HRS 2000 data collection. Across four strata, respondents were assigned to \$20, \$30 or \$50 treatment groups. The findings were:

- Those given \$50 had consistently higher response rates than the \$20 group.
- The response rate for those given \$30 generally (but not always) fell between the \$20 and \$50 groups.
- Providing the \$50 incentive reduced the number of face to face contacts by .115 and number of telephone contacts by .646 compared to those receiving \$20. The number of interviewer hours were also substantially reduced.

As a result, the HRS unconditional incentive was increased to \$40 in 2004, and increased again to \$50 in 2006 for a sub-sample of respondents who did a face to face interview.

#### *British Household Panel*

In 2004 (Wave 14) a split-sample experiment was run to test the effect of increasing the value of the adult interview incentive from £7 to £10, and increasing the value of the young person's incentive (aged 11-15) from £4 to £5. Half of those eligible were randomly assigned to receive the lower (standard) amount, while the other half received the higher amount. The findings were as follows:

- Those adults in the £10 incentive group had significantly higher response rates.

- The increased incentive appeared to have a greater effect on response for those who were eligible but not interviewed at the prior Wave (Wave 13). This suggests that increasing the incentive may be an effective strategy for prior Wave refusals.
- A significantly higher proportion of respondents who were interviewed by telephone at the prior Wave (due to face to face refusal), were converted to a full interview at Wave 14 if they received £10 in comparison to £7.
- The effect of the increased incentive was greatest amongst those aged 16-24 who were the existing group with the lowest response rate.
- For new 16 year olds who did a youth self completion last Wave (when aged 15) and were now eligible for the adult interview - the £10 incentive had a bigger effect on participation than the £7 incentive.

Examples of studies that have targeted specific groups

#### *Health and Retirement Study*

The Health and Retirement Study (HRS) invested much resource into targeting non-respondents at the end of its first Wave. As part of a non-response study, a second phase sub-sample of reluctant respondents was drawn (Juster & Suzman, 1995). Three main experimental groups were selected: 1) those deemed as 'final refusals' after many refusal conversion attempts; 2) those classified as 'reluctant cases' that had had fewer contacts than the first group; and 3) all remaining non-respondents. All groups received a standard letter but some were offered \$50 to participate while others were offered \$100. The non-response study successfully increased the overall HRS response rate by 4 percentage points.

Lengacher et al (1995) reviewed the possible effects of such large Wave 1 incentives on participation at Wave 2. They concluded that combating non-response at the initial Wave through the introduction of large incentives to refusers seemed to have no long term effect on participation at the second Wave. It seems that they did not expect to get the same incentive amount again.

Lengacher et al went on to establish that enjoyment of the HRS interview was found to play a role in influencing Wave 2 behaviour (namely the decision to participate again) except for those who had received the large incentive. It may be the case that their memory of the large incentive diminished the salience of the Wave 1 interview, but they perceived the large payment at the first Wave to cover their participation in later interviews.

#### *The Survey of Income and Program Participation (SIPP)*

At Wave 7 of SIPP a booster incentive of \$20 was given to all low income households that had received a Wave 1 incentive. Up to then, no other incentives had been used since Wave 1. This strategy resulted in a reduced non-response rate at Wave 7.

At Wave 8 and 9 an experiment was set-up to test the effect of prepaid monetary incentives to non-responding households at the prior Wave (Martin et al, 2001).

Assignment to treatment groups was independent of the incentive group they were in at Wave 1. In this experiment they received either \$20, \$40 or no incentive. The results were as follows:

- Both the \$20 and \$40 groups obtained significantly higher conversion rates than the control group, but the \$20 and \$40 groups did not differ significantly.
- \$20 yielded improvement in the high poverty stratum with no further gain from \$40.
- In contrast, only the \$40 had an effect in the low poverty stratum.
- Incentives were more effective for conversion of refusals than any other type of non-interview.
- College educated households were less likely to be converted to interviews.

This experiment has shown that payment amounts may have differential effects on groups of individuals, and that it may be possible to tailor incentives to attract longitudinal members back to the survey.

#### *National Longitudinal Study of Youth*

The US National Longitudinal Study of Youth started in 1979 and offered an incentive of \$10, until 1996 when it increased to \$20.

In 2002, the base rate for a face to face interview rose to \$40. However, they attempted to get more cooperative 'easy' cases at lower cost by asking respondents to call in to do the interview by telephone. If they chose to be interviewed by phone they were offered a higher incentive amount – either \$60 or \$80. The main findings were:

- Response rates were higher, and the costs per interview lower, if respondents were given the opportunity to phone in.
- The \$20 difference in incentive for telephone cases had no significant effect even for reluctant respondents.

Overall, increasing the incentive amongst those eligible for a telephone interview had the advantage of increasing response rate and reducing costs per interview.

## Appendix L Parental Interviews

### Who was interviewed at each Wave of the first LSYPE cohort

The Wave 1 interview consisted of five modules. The sample member completed one module – the young person interview, which lasted approximately 35 minutes. Adult interviews were also completed for household information, including the main parent interview, second parent interview<sup>20</sup> and child history. These lasted approximately 55 minutes altogether. The total interview time was 1 hour and 30 minutes.

For Wave 2 the young person interview lasted approximately 35 minutes. There was considerable variation in the length of adult interviews depending on whether the adult was interviewed at Wave 1 or not. Interviews with those not interviewed at Wave 1 took longer as interviewers had to collect additional data missed at Wave 1. The total target interview time was 1 hour and 10 minutes.

The main parent was defined as the parent most involved in the young person's education, and the second parent was defined as adults other than the main parent who had a parental relationship to the young person (i.e. a natural, step, adoptive or foster parent). During the Wave 2 fieldwork it was necessary to change the definition of a second parent as it became apparent that the relationship of 'partners' of the main parent were not necessarily identified as being a 'parent' or guardian to the young person. Therefore the definition of the second parent changed and all second parents were identified as those who were a partner or spouse of the main parent. Due to the number of second adults not interviewed at Wave 1, considerable efforts were also made by interviewers to speak to these people at Wave 2.

It must also be noted that the main parent and second parent roles at Wave 1 were not always carried forward to Wave 2, even in cases where the parent(s) were still living with the child. For example, at Wave 1 the mother may have answered the main parent questionnaire and the father the second parent, but at Wave 2 these roles could have reversed. It is however possible to compare the positions in the household grid at Waves 1 and 2 to identify cases where this happened.

Wave 3 consisted of four modules. The sample member completed one module which lasted approximately 20 minutes. Adult interviews were also completed for household information, the main parent interview and individual parent interviews. Unlike at previous Waves, in households with 2 parents, there was no second parent interview. The main parent answered the individual parent questions on behalf of both parents. These lasted approximately 15 minutes altogether so considerable time (and cost) was saved. The total target interview time was 35 minutes. It was also possible that by Wave 3 some of the young persons no longer lived with their parents.

---

<sup>20</sup> In 14 cases at Wave 1, no second parent was identified in the household interview but a second parent interview was subsequently conducted. None of these households responded at subsequent Waves.

At Wave 4, the parental module was completed by the main parent. This comprised of three parts: a main parent section (asked primarily of boost respondents, with a few questions asked of all respondents), and two individual parent sections, the first relating to the main parent and the second to the main parent's partner (if applicable). The individual parent sections collected details about the employment, education and training, and health of each parent. If the partner was present at the time of the interview, the individual parent partner questions were asked directly of them. If not, then the main parent was asked to answer on behalf of their partner. Overall the parental module lasted approximately 10 minutes for main sample parents, and 25 minutes for minority ethnic boost sample members.

Identification of the main parent differed between the main sample households and the minority ethnic boost sample households. In the main sample, the main parent interview was conducted with either parent/guardian. Interviewers were asked to pick the parent who they felt was more likely to take part. This is a change from previous years where the main parent was identified as the parent most involved in the young person's education. This was no longer necessary at Wave 4 as the parent interview focused largely on the employment status and health of the parents themselves. In the minority ethnic group sample, the main parent was determined by the following order of priority: 1. Natural mother; 2. Natural Father; 3. The parent most involved in the sample member's education.

The content is summarised in Appendix B, Table B.1.

## **Comparative longitudinal studies that include a parental interview**

### **Millennium Cohort Study**

The Millennium Cohort study, like the first LSYPE cohort, distinguishes between one parent who is the 'main respondent' and the other parent who is the 'partner respondent'. The 'main respondent' completes a full interview covering all topics, while the 'partner respondent' is asked a subset of questions from the Main Interview. The Millennium Cohort study also includes an option to do a Proxy Partner Interview for those who were away for the entire fieldwork period or who were incapable of completing an interview themselves.

At MCS1, there was a Main Interview in 18,532 of the 18,552 families. There was someone eligible for a Partner Interview in 15,358 families and an interview was completed in 13,225 of these cases. Proxy data were collected on 216 partners (of the 235 who were eligible) but interview data are completely missing for 1,917 two 'parent' families. The vast majority of the main respondents were female. Since the main respondent was asked questions about pregnancy and delivery the presumption was that, wherever possible, the natural mother should be the main informant. Some of the cases where roles were reversed were because of language problems.

At MCS2, there was someone eligible for the Main Interview in 15,588 of the 15,590 productive families and an interview was completed in 15,448 cases. There was

someone eligible for a Partner Interview in 12,856 families and an interview was completed in 10,479 of these cases, with data by proxy in 233. There are 2,154 two 'parent' families with data missing on the partner, and 63 with data missing from the main. There are also 79 families with some data (e.g. child assessments) but no interview data from either a main or a partner respondent. The Main respondents are again overwhelmingly female, but the number of them who are not natural mothers had increased. Part of this change is an increase of lone-father informants but it is mostly due to a rise in the number of two-parent families where the main response was collected from the father (97% of the partners were natural fathers).

At MCS3, a Main Interview was conducted in 15,210 of the 152,246 families. There was someone eligible for a Partner Interview in 12,189 families and an interview was completed in 10,475 cases, with proxy data collected in a further 287. Information was not collected on partners in 1,408 couples, and from main respondents in 19 families where the partner responded. In 36 cases there were no interviews in the dataset from any parent. The proportion of main informants who were natural mothers again dropped, to 97 percent (14,792). The number of female main respondents who were not natural mothers hardly changed from MCS2 (58). It was the numbers of main respondents who were men that also changed. The number of natural fathers completing the main interview was 394 (more than double the 185 at the age 3 survey). Seventy two were lone fathers and the rest were part of a couple. The switch to a male informant would have arisen in cases where the natural mother no longer lived with the child and the father, and where the father elected to be treated as the main carer.

In the vast majority of cases at all sweeps the natural mother did the main interview and the natural father the partner interview.

### Growing Up in Scotland

The Growing Up in Scotland Study had an interview with the child's carer in the first sweep, and then invited a second parent to be interviewed at a second sweep.

Interviews were carried out in respondents' homes, by trained social survey interviewers using laptop computers (otherwise known as CAPI - Computer Assisted Personal Interviewing). The interview was quantitative and consisted almost entirely of closed questions. There was a brief, self completion section in the interview in which the respondent, using the laptop, inputted their responses directly into the questionnaire programme. Interviews were conducted with the child's main carer in the first sweep. At this sweep, primarily because of the inclusion of questions on the mother's pregnancy and birth of the sample child, interviewers were instructed as far as possible to undertake the interview with the child's mother. At sweep 2, interviewers were instructed to undertake the interview with the sweep 1 respondent. Where this was not possible or appropriate, interviews were conducted with the child's main carer. In practice, most interviews were undertaken with the sweep 1 respondent and this was usually the child's mother.

### The National Child Development Study (NCDS)

NCDS is a continuing, multi-disciplinary longitudinal study which takes as its subjects all the people born in one week in England, Scotland and Wales in March 1958. NCDS has its origins in the Perinatal Mortality Survey. Sponsored by the National Birthday Trust Fund, it was designed to examine the social and obstetric factors associated with stillbirth and death in early infancy among the children born in Great Britain in that one week. Information was gathered from almost 17,500 babies. For the birth survey, information was obtained from the mother and from medical records by the midwife. For the purposes of the first three NCDS surveys, information was obtained from parents (who were interviewed by health visitors), head teachers and class teachers (who completed questionnaires), the schools health service (who carried out medical examinations) and the subjects themselves (who completed tests of ability and, latterly, questionnaires).

At each of the first three follow-ups, information was obtained from four main sources: the children themselves, the parents, Local Authority Medical Officers, and schools. At each age, the parents (in fact most commonly the mother alone) were interviewed in the home by an officer of the local authority, usually a Local Authority Health Visitor, using a structured interview schedule.

### The 1970 British Cohort Study (BCS70)

BCS70 is a continuing, multi-disciplinary longitudinal study which takes as its subjects all those living in England, Scotland and Wales who were born in one particular week in April 1970. Data has been collected from a number of different sources, and in a variety of ways. In the birth survey, information was collected by means of a questionnaire that was completed by the midwife present at the birth, and supplementary information was obtained from clinical records.

The five-year and ten-year surveys were carried out by the Department of Child Health, Bristol University and the survey at these times was named the Child Health and Education Study (CHES). In 1975 and 1980, parents of the cohort members were interviewed by Health Visitors, and information was gathered from head and class teachers (who completed questionnaires), the school health service (which carried out medical examinations on each child), and the subjects themselves (who undertook tests of ability). Overall, interviewers reported that parents were very positive about the experience of taking part in the study. As expected, given the nature and source of the sample, there were few problems gaining cooperation. However, fathers living with their baby were harder than mothers to contact, though in most cases interviewers felt that participation could ultimately have been achieved. Partly in view of this, the flexibility to conduct the father interview first was valued by interviewers. This was taken forward into the computerised instrumentation.

### The Avon Longitudinal Study of Parents and Children (ALSPAC)

ALSPAC recruited more than 14,000 pregnant women with estimated dates of delivery between April 1991 and December 1992. These women, the children arising

from the index pregnancy and the women's partners have been followed up since then and detailed data collected throughout childhood.

#### Growing Up in Australia (LSAC)

The Longitudinal Study of Australian Children (LSAC) is designed to identify policy opportunities for improving support for children and their families and for early intervention and prevention strategies. The study commenced in 2004 and will continue to follow the progress of around 10,000 children until at least 2010. Families were recruited and interviewed in 2004, with interviews conducted every 2 years for 4 Waves of data collection (2004, 2006, 2008, 2010).

A home interview is conducted at each Wave (120 minutes in the home), self completion questionnaires for Parent 1 and 2, direct child assessment, time use diary, interviewer observations, carer/teacher questionnaires and data linkage. Parent 1 is defined as the parent who knew the child best (97% mothers). The home interview with parent 1 was followed by a self completion questionnaire (which could be collected by the interviewer or sent back by the respondent, and had an 85% response rate). Parent 2 was defined as Parent 1's resident partner and were almost all fathers. They had a self completion questionnaire only (which could be collected by the interviewer or sent back by the respondent, and had a 78% response rate).

#### The National Longitudinal Study of Adolescent Health

The National Longitudinal Study of Adolescent Health (Add Health) is a longitudinal study of a nationally representative sample of adolescents in grades 7-12 in the United States during the 1994-95 school year. Data was collected from the adolescents themselves, and from their parents, siblings, friends, romantic partners, and school administrators.

In the first Wave, the in-school survey was carried out in 140 schools and was followed up a year later by an in-home interview of the study youth and the 'principal care giver' (typically the mother). The care giver was asked to complete an interviewer-assisted questionnaire. The adolescents were interviewed once again in the home a year later, in 1996. Approximately 10,000 adolescents participated in all three Waves.

# Appendix M      Communication plan

## **Workshops with young people**

Four one hour workshops were held with groups of 13-15 year olds young people in four schools of varying levels of achievement in the South East:

- **Group 1:** 10 boys (5x13yr olds and 5x 15yr olds) in a single sex, average achieving comprehensive school in North London.
- **Group 2:** 12 girls aged 13-15 in a private, high achieving all girls school.
- **Group 3:** Mixed group of 13-15yr olds in an average achieving independent school in Milton Keynes.
- **Group 4:** Mixed group of 13-14 year olds in a low achieving state school in Arnos Grove.

## **Detailed communications materials**

**The website:** The website should become the study 'hub' over time with respondents viewing the latest findings and seeing where the study has appeared in the media. The existing website 'LSYPE' is good in that it is informative and open about the study's purpose. However, it is too wordy (young people just don't read text online) and uninteresting. A new site should be developed which gives young people (and their parents) the opportunity to:

- Find out about the study and why it's so important
- See how it has helped improve people's lives
- Sign up for email updates and text alerts
- Read about latest study findings
- Create an 'account' where they can update their personal details.

**Information pack for young people and parents:** The pack could contain a range of materials designed to answer questions about the study and inspire people to take part:

- Booklet. Evidence from other surveys suggests that an attractive, engaging and aspirational booklet reinforcing the purpose, importance and uniqueness of the study can create a positive perception of the study in the minds of the respondent and encourage them to take part. Interviewers can also have extra copies to hand out on the doorstep.
- Membership card. To allow young people to register online, update their details and claim their incentives.
- Parental letter of consent.

**Emails and text updates:** Short, text based email updates should be sent to young people regularly to keep them up to date with the study and drive traffic back to the website. These emails can be plain text (i.e. they do not need to use HTML

templates) as there is growing evidence to suggest that people open plain, personal emails more often than graphically designed newsletters. Similarly, there would be value in using SMS alerts to ask young people to visit the website to update their emails or to find out the latest about the study. Interviewers should request young people's relevant contact details and permissions during the initial interview. It is also important that young people are given the opportunity to opt out of this feature.

## **Methodology and detailed insights from research among research users**

### **Research to inform this strategy**

NatCen undertook 12 semi-structured face to face and telephone interviews with current and former members of the research team at DfE with responsibility for LSYPE, senior staff within DfE, two of the three survey organisations who form part of the consortium who undertake the survey (and have done so since it first started), LSYPE consortium members from other Government Departments, policy makers within DfE, government and non-government data analysts. The interviews covered a range of issues, with topic guides being tailored to reflect the different types of stakeholder mentioned above.

Approximately 20 past and potential research users were also consulted at a User Group Seminar day in May. A questionnaire was circulated to explore what tools people use in the workplace to find out and access research, and in what format they would like to receive findings.

## **Insights and what this means for communication**

### **Awareness and knowledge of LSYPE and the value of longitudinal studies is patchy**

- Levels of awareness of LSYPE were found to be variable among those interviewed – both within and outside DfE. A perennial problem is the lack of awareness/understanding of what a longitudinal study like LSYPE has to offer. This lack of understanding was particularly acute among policy makers, who tended to see LSYPE as providing cross-section data on a current issue (how many people did x? how does this proportion change over time?) rather than providing answers to fundamentally longitudinal questions (how do educational trajectories differ and why? what factors have the most influence on the education outcomes of individuals?).
- Policy makers, with one exception, did not realise that LSYPE collected information from parents.
- The distinction between LSYPE and YCS was not always clear to policy makers and this appeared to stem from the fact that the results of both studies are included in one publication (though the LSYPE also has stand alone publications).
- Evidence suggests that Departments and divisions within Departments tend to have a very limited knowledge of what others are doing with the data.

### **Data and report need to provide findings in a range of ways**

- Policy colleagues left it to analysts to interrogate the data and had no appetite for additional bespoke packages of data or reports. They would prefer more face to face briefings/presentations as this is the communications medium that they use/ most prefer.
- Nearly all User Group Seminar respondents proactively looked for research themselves (rather than being supplied with research findings by peers). When asked what tools they use to perform their work, they cited a mix of trade and specialist publications, internal newsletters, external websites (e.g. ESDS), intranets and e-newsletters. When asked how best they wanted to hear about research findings they specified a mix of face to face (informal chats and briefings), emails, publications and websites. So there is a need to supply data/findings in a range of formats, using a number of channels to ensure they are catering to everyone's needs and maximizing the number of people interacting with the study.
- Some User Group Seminar respondents felt that data outputs and datasets should be marketed to make them easier to pick up and engage with. However, it was recognised that this is published from a Government Department and that external communications are cleared through Departmental press offices, which reduces the independence Departments have in publicising data.

### **Findings need to be shareable**

- Nearly all User Group Seminar respondents reported that they share research with colleagues via word of mouth or round robin emails to colleagues with links attached. Many also attended seminars and conferences, so it would be helpful to supply findings in formats that support this activity e.g. PowerPoint presentations, emails with short PDF/word report summaries and links to the website, detailed reports.

### **Data need to be easy to use**

- Some User Group Seminar respondents felt that the LSYPE data sources are complex and whilst the data are well documented, they are still very difficult to use, simply because of the size of the datasets and the number of variables they contain. They suggested that the data could be simplified or broken up into smaller datasets focusing on specific areas (although there was recognition that this requires resource to do). There was a view that this may encourage PHD students or inexperienced users to use the data, with the aim that they would ultimately become expert over time.
- A subgroup of User Group Seminar respondents unanimously said that the data would be easier to use in future with more derived variables, particularly for the life history file which is very difficult to use. Again it was recognised that this takes time and resource.

### **Clarity around the study brand**

- Some User Group Seminar respondents were confused about the two names for the study. LSYPE is generally used to describe the study within the Department and amongst research users, and Next Steps is what the respondents know the study as. Some did not know that the two names relate to the same study. Further consideration of the pros and cons of having separate brands for young people and research users is therefore advised, not least because it raises questions of how the study should be positioned in media read by both sets of audiences.

**Ref: DFE-RR048**

**ISBN: 978-1-84775-810-1**

**© NatCen**

October 2010