Enhancing data around early life in Understanding Society: scientific opportunities and considerations

Michaela Benzeval (University of Essex)
Non-Technical Summary

*Understanding Society* already provides unique insights into children’s lives in the 21st century and will continue to track them as they develop into adults. The Study naturally follows the development of families over time from couples forming to babies being planned, conceived and born, to children growing up in their family context (whether both parents remain in the house or not) and transitioning to adulthood. The Study also enables research that examines childhood in different contexts, whether these are geographical, ethnic group and generation status, family structure, economic circumstances, cohort or period. It also allows for the investigation of inter-generations effects whether they be those of parents and grandparent on children, or of children on their parents.

In this think piece we have outlined ten areas where we are currently working, or have plans, to improve the multi-dimensional biological and social data that *Understanding Society* provides on families across households and generations to significantly enhance the data we produce for scientific and policy research at this crucial life-stage.
Enhancing data around early life in Understanding Society: scientific opportunities and considerations.

Professor Michaela Benzeval (University of Essex)

Abstract Understanding Society already provides unique insights into children’s lives in the 21st century and will continue to track them as they develop into adults. The Study naturally follows the development of families over time from couples forming to babies being planned, conceived and born, to children growing up in their family context (whether both parents remain in the house or not) and transitioning to adulthood. The Study also enables research that examines childhood in different contexts, whether these are geographical, ethnic group and generation status, family structure, economic circumstances, cohort or period. It also allows for the investigation of intergenerations effects whether they be those of parents and grandparent on children, or of children on their parents.

In this think piece we have outlined ten areas where we are currently working, or have plans, to improve the multi-dimensional biological and social data that Understanding Society provides on families across households and generations to significantly enhance the data we produce for scientific and policy research at this crucial life-stage.

Keywords: birth cohort, event histories, children, fertility, life course studies

Corresponding author: Michaela Benzeval, Institute for Social and Economic Research, University of Essex, Wivenhoe Park, Colchester, CO3 4SQ. mbenzeval@essex.ac.uk
## Contents

Summary .......................................................................................................................... 4

Introduction ................................................................................................................... 6

What scientific questions can *Understanding Society* address around childhood experiences in 21st Century? ......................................................................................... 7

Families beyond households........................................................................................ 10

What scientific questions can *Understanding Society* address about the experiences of different birth cohorts in the UK? ........................................................................... 13

What are the new research opportunities and scope for expanding *Understanding Society* data on early life? ............................................................................................ 15

Women of childbearing age and new mothers in *Understanding Society* ............. 16

Sample sizes ................................................................................................................. 16

How do *Understanding Society* mothers compare to the general population? ....... 17

New babies in *Understanding Society* ........................................................................ 19

Next Steps: Developing the *Understanding Society* pregnancy, early life and family data .............................................................................................................................. 20

Regular biological data collection .............................................................................. 21

Event triggered data collection .................................................................................... 21

Data collection in pregnancy and the first year of life .................................................. 22

References ................................................................................................................... 24
Summary

*Understanding Society* already provides unique insights into children’s lives in the 21st century and will continue to track them as they develop into adults. The Study naturally follows the development of families over time from couples forming to babies being planned, conceived and born, to children growing up in their family context (whether both parents remain in the house or not) and transitioning to adulthood. The Study also enables research that examines childhood in different contexts, whether these are geographical, ethnic group and generation status, family structure, economic circumstances, cohort or period. It also allows for the investigation of inter-generations effects whether they be those of parents and grandparent on children, or of children on their parents.

In this think piece we have outlined ten areas where we are currently working, or have plans, to improve the multi-dimensional biological and social data that *Understanding Society* provides on families across households and generations to significantly enhance the data we produce for scientific and policy research at this crucial life-stage.

**Action 1**: The Study team to develop a family group identifier and share with data users to enable inter and intra generation research across households.

**Action 2**: The Study team to draw together lessons from research on how to improve data on separated and/or non-resident co-parenting families, discuss with experts in field, and implement changes in the main study where possible.

**Action 3**: Additionally as part of our Wave 13+ bid (not yet funded), we included plans to expand the survey at Wave 14 to include the ‘other parts’ of separated/non-resident co-parenting families not currently included in the sample. This might include people who have a child with an *Understanding Society* sample member without living together, or who are temporary sample members (TSMs) or the ex-partners and children of parents who move into an Understanding Society household, If the bid is successful, the Study team will set up a working group to advise on the design and piloting of the separated families extension to the Study in preparation for Wave 14.

**Action 4**: The Study team plan to undertake some demonstration projects using data on pregnancy outcomes and child development from age 1 to 8, to highlight their availability.

**Action 5**: Drawing on external expertise through workshops and other consultative mechanisms, the Study Team plan to enhance the questions we ask at these ages, and expand coverage where appropriate – e.g. to ask men about their babies born outside the household as well as about those within it, to increase the age of women we ask fertility and child birth questions - to improve the value of the Study for early life research.

**Action 6**: The Study team plan to examine and revise our retrospective histories data collection strategy to ensure we have an effective system for collecting missing histories from participants.

**Action 7**: The Study team will maximise the number of *Understanding Society* babies that are included in the sample by improving our identification of newborns, developing approaches to ensure we collect timely pregnancy and early life data on as full a sample as possible, for example with shorter questionnaires, asking questions about the first year of life to fathers as well as mothers, changing the proxy interview to include these questions, and by
collecting data retrospectively in missed in the first years of life. In addition we will reconsider our following rules for newborns born outside the household and/or to TSMs.

**Action 8:** The Study team will maximise learning from IP12 so that the inclusion of biomarkers can be rolled out to adults and children in *Understanding Society* going forward. Once processes for effectively measuring the IP biomarkers are established, we will investigate broadening the accurate participant-led measurement of biomarkers to other measures, and expanding the range of measures we collect in childhood.

**Action 9:** The Study team will continue to experiment with protocols for event triggered data collection so that the most effective approach can be rolled out (if funded) in 2021. To develop appropriate modules to be asked either in the main study or through event triggered data collection about parents-to-be’s expectations, psychological wellbeing and behaviours during pregnancy and in the early months of the baby’s life.

**Action 10:** Drawing on key lessons from this first pregnancy feasibility study, the Study team will design more effective materials to explain the value of data collection at this life-stage to parents, and develop protocols for data collection that can be trialled that minimise burden and invasiveness while maximising the key data that can be collected.
Introduction

The 2017 Longitudinal Studies Review recommended that ESRC commissions a new UK population-representative birth cohort (Davis-Keen, 2017), and subsequently ESRC have asked what role data from Understanding Society might play in meeting these research needs. The Review identified a number of different requirements that a new birth cohort might fulfil:

- address the current data gap on UK children since 2000/01;
- maintain a continuous flow of data to track societal changes and their impacts on generations born in the 21st century, and to address policy questions related to these changes;
- concentrate not only on children, but also could collect data on resident and non-resident fathers and mothers in order to adequately address the gene-environment interplay, to better understand the dynamics of separated families, and to enable more research on the intergenerational transmission of inequalities;
- enable research on early childhood’s prediction of lifespan effects including the interplay not only of post-natal, but also pre-natal factors and exposures, as well as environmental factors (including pollution and climate change) might be of great research interest;
- be interdisciplinary and innovative and include the collection of biomarkers and the study of genetic, epigenetic and other omics effects to map children’s early and later exposures;
- to be simpler and hence less risky than the Life Study.

Cohort studies provide rich data on the development of the population at key life-stages and the impact of circumstances and experiences at one life-stage on the next. Having cohorts born each decade with similar data has enabled researchers to compare childhood experiences and their impact on later life over time. However, birth cohort studies are not the only kind of study that can provide insights into children’s lives, within their family and societal context, as they develop.

Understanding Society, as a household panel study, collects data across a wide range of domains on people’s everyday lives, which means it can shed light on the impact of contemporaneous societal, political, economic and policy changes on people at all life-stages at the same time. It also tracks people over time so that researchers can investigate how circumstances earlier in life, at different life-stages, can influence later behaviours, attitudes and outcomes. It follows people in their original households and as they move on to form households of their own, so it can investigate linked lives within and across generations. Given this, it provides invaluable insights into children’s lives in the 21st century, within their family context, and with ongoing funding, will continue to track the children, and their families, as they age, in the future.

The purpose of this think piece is therefore threefold. First to identify the kinds of scientific questions that can already be answered by Understanding Society, and identify modest enhancements, some already in train, that will expand these research opportunities. Second, we identify the new research opportunities and scope for expanding Understanding Society data on early life by investigating potential samples sizes of new babies and their family characteristics. Finally, we set out our initial thoughts for how we might enhance data collection at this crucial life-stage.
What scientific questions can *Understanding Society* address around childhood experiences in 21st Century?

As a household panel study, *Understanding Society* (University of Essex, 2018) naturally follows the development of families over time from couples forming to babies being planned, conceived and born, to children growing up in their family context (whether both parents remain in the house or not) and transitioning to adulthood. While it captures biological children from (before) birth, it also allows for the inclusion of other children as they join the household - for example both adopted and foster children (Brown *et al.*, 2019) – as well as step children as families separate and reform. As all people in the household are interviewed, and followed as they move out of their original household into new ones, the Study effectively enables intra and inter-generational research within and across households. The Study also enables research that examines childhood in different contexts, whether these are geographical, ethnic group and generation status, family structure, economic circumstances, cohort or period. Box 1 provides some brief illustrations of the kinds of research that has been done on childhood, and its effects on adult lives, in *Understanding Society*. 
Before children: Fertility intentions
Research using *Understanding Society* (and its predecessor the British Household Panel Survey) has examined the changing patterns of fertility intentions and outcomes (Berrington, 2004; Iacovou and Taveres, 2011); labour market participation and fertility intentions (Zhou and Kan, 2019), and gender role attitudes and fertility plans (Okun and Raz-Yurovich, 2019).

Other researchers have compared fertility in UK internationally, and/or among different ethnic groups (Kulu *et al*., 2017; Wilson, 2018). It also enables the examination of age, period and cohort effects, as a result of collecting the same data repeatedly over time and across all age cohorts. Research on cohorts is described in Box 2; an example of research on period effects with *Understanding Society* is the impact of the great recession on fertility (Lyons-Amos and Schoon, 2018).

Family life and childhood
*Understanding Society* collects data on children from birth to 10 from parents (see below) and currently begins to interview children at age 10, and a number of research themes have emerged from the data to date. For example, there is a range of research on child behaviours and outcomes — travel, wellbeing, education (Wu *et al*., 2019, Orben *et al*., 2019, Brown *et al*., 2019) — as well as the development of attitudes, aspirations and expectations (Berrington *et al*., 2016; Martin and Mellon, 2018; Fox *et al*., 2019). *Understanding Society* also enables intra family relations to be explored. For example, research has examined the role of siblings on the educational aspirations and wellbeing of each other (Yucel *et al*., 2015; Bu, 2016; Tippett and Wolke, 2015), inter-generational transition of behaviours (McAloney *et al*., 2014), the impact of parent and children’s mental health on each other (Webb *et al*., 2017), and the effect of couple’s relations with each other on their involvement with their children (Poole *et al*., 2014). It also provides evidence on family circumstances during childhood (Green *et al*., 2018), and the experiences of growing up in separated families for the children and parents (Brewer *et al*., 2016; Poole *et al*., 2016; Haux *et al*., 2017, Haux and Luthra, 2019). As well as within household family relations, *Understanding Society* has also been employed to examine the role of grandparents on children’s lives (Zhang and Li, 2019).
Youth Adult transition

*Understanding Society* enables the study of the transition from childhood to adulthood across a range of dimensions, and the exploration of family circumstances pre-conception and birth, in childhood and later life attitudes, behaviours and outcomes. For example, by following adolescents as they move into adulthood, the Study has been employed to examine how transitions to different adult roles vary by ethnic group and family background (Berrington, 2018), as well the inter-relations between different kinds of adult transitions such as leaving education and looking for employment (Schoon and Lyons-Amos, 2016; Zuccotti and O'Reilly, 2019), leaving the parental home (Bayrakdar and Coulter, 2018), residential moves (Pelikh and Kulu, 2018) and becoming a home owner (Bayrakdar *et al*., 2019). It has also been employed to examine how unhealthy behaviours (smoking) are formed over the youth-adult transition and how these are impacted on by changing tobacco control policies over time (Green *et al*., 2016).

The influence of childhood on later adult attitudes, behaviours and outcomes

Researchers have also employed the Study to examine the impact of circumstances and the environment and attitudes during childhood on later life circumstances and behaviours. For example, a number of studies have employed retrospective reports (Karimi *et al*., 2019; Hughes *et al*., 2019) and data measured during childhood (The Children's Society and Barnados, 2018; Booker, under review) to examine childhood circumstances and adult health outcomes. Others have looked at where people lived as children and the impact on later life health. For example, one study examined migration experiences of children and their parents during childhood and the child’s wellbeing as an adult (Eremenko and Bennett, 2018). Other studies have examined more structural issues such as social mobility (Zwysen, 2015), or whether parental attitudes during childhood influence different kinds of behaviours and attitudes as adults, for example, savings behaviours (Brown and Taylor, 2016), political values (Fox *et al*., 2019) or gender attitudes (Platt and Polavieja, 2016).

The impact of having children on parental outcomes and behaviours

Much research has been carried out, including with *Understanding Society*, on mother’s return to work decisions (e.g. Chung and van der Horst, 2018a; Khoudja and Platt, 2018) and ongoing working patterns (e.g. Chung and van der Horst, 2018b), child care arrangements (e.g. Bordone *et al*., 2018) and the impact of becoming a parent on attitudes and behaviours (e.g. Zhou, 2017; Thomas *et al*., 2018). However, little of this research uses information on how the characteristics and behaviour of the child(ren) may influence such decisions and behaviours. With the data that already exists in *Understanding Society* on young children’s development and behaviours (see Box 2), it would be possible to enrich this field by exploring how differences in children’s sleeping, feeding, development influence parental employment and child care decisions, as well as their wellbeing, sleep and health. Some research like this already exist at older ages, for example, we found several papers on how children’s health influences their parents’ wellbeing (Webb *et al*., 2017; Booker and Sacker, 2011).
Families beyond households

We are currently developing resources to examine family relations beyond households in a number of ways that will create exciting opportunities for really novel research.

Firstly, as the Study continues, opportunities to investigate these multi-generational family exchanges will grow. We are currently developing an easier way for data users to identify multi-generational families across households within the Study. Although still in development, this approach has enabled us to identify 2,699 three-generation-generation generational families (grandparents, parents, children), which include 6,714 people. We will also identify wider family connections such as aunts/uncles and cousins, so the role of wider family networks in childhood development can be explored.

**Action 1:** The Study team to develop a family group identifier and share with data users to enable inter and intra generation research across households.

Secondly, one specific benefit of *Understanding Society* for research in childhood is that the study design means when couples separate, in principle, both halves are followed into new homes and continue to be followed as they form new relationships. Indeed *Understanding Society* has been identified by an independent review as the key study for understanding the impact of parental separation on childhood (Bryson *et al.*, 2017). However, our data are currently sometimes limited by our following rules when couples separate, for example, we currently only follow original or permanent samples members when they leave the household (see Fumagalli, 2019 for detailed discussion of temporary sample members (TSMs)), the attrition of the parent leaving the family home from the survey, which can be high (Bryson *et al.*, 2017), and by the questions that are asked (Bryson and McKay, 2018, Haux and Luthra, 2019). We have undertaken qualitative research to identify ways of improving our ability to track separating parents when they leave the family home (Kantar Public, 2017) and carried out experiments on the way in which questions around separation and non-resident children are asked (Bryson and McKay, 2018). In another project being carried out concurrently, colleagues have conducted interviews with both halves of separated couples to understand issues around shared care and how our questions on co-parenting can be improved in future (Haux and Luthra, 2019).

**Action 2:** The Study team to draw together lessons from research on how to improve data on separated and/or non-resident co-parenting families, discuss with experts in field, and implement changes in the main study where possible.

**Action 3:** Additionally as part of our Wave 13+ bid (not yet funded), we included plans to expand the survey at Wave 14 to include the ‘other parts’ of separated/non-resident co-parenting families not currently included in the sample. This might include people who have a child with an Understanding Society sample member without ever living together, or who are TSMs, or the ex-partners and children of parents who move into an Understanding Society household. If the bid is successful, the Study team will set up a working group to advise on the design and piloting of the separated families extension to the Study in preparation for Wave 14.

Surprisingly, given data are available on pregnancy outcomes and child development from age 1 to 8, we could find no research papers using these data. The data available in *Understanding Society* at these ages is highlighted in Box 2. There are a range of reasons why these data might not have been used. First, the availability of these data are not well
known, indeed a presenter at the 2019 *Understanding Society* Scientific Conference in July confidently told the audience the Study had no data on childhood. Obviously there is a circularity problem, if researchers are not aware of the data they do not use them, and if no one is using the data and publishing with them, then no one will be aware of them.

Secondly, compared to the cohort studies, *Understanding Society* currently has more modest data about child outcomes and behaviours at these ages, and in the past some of the parent-child research questions that might be addressed were hampered by parenting questions being asked about children in general rather than for each specific child (this has now been changed). By design the birth cohorts have much richer information at these life-stages. However, as illustrated in Box 1, *Understanding Society*, could provide some valuable insights drawing on these data, given other unique features of the Study, such as having data on the whole household, the ethnic minority boost sample, providing continuous evidence on children moving through these ages, across different birth cohorts, etc.

Given the above, in the Wave 12 (now successful) grant application we included a research post to promote and improve the child development data.

**Action 4:** The Study team plan to undertake some demonstration projects using data on pregnancy outcomes and child development from age 1 to 8, to highlight their availability.

**Action 5:** Drawing on external expertise through workshops and other consultative mechanisms, the Study Team plan to enhance the questions we ask at these ages, and expand coverage where appropriate – eg to ask men about their babies born outside the household as well as about those within it, to increase the age of women we ask fertility and child birth questions - to improve the value of the Study for early life research.
Box 2 Pregnancy and child development question under age 8

Pregnancy and first year
Men aged 15-64 asked: if fathered a child since the last interview (whether within or outside household)
Women aged 15-49 asked about pregnancies since last interview (loop of questions for up to 4 pregnancies), type of conception, outcome of pregnancy. For live births – type of delivery, whether smoked or drank in different pregnancy trimesters.
For each live birth (whether with or outside household; mother or father or both) – date of birth, sex of child, whether in household, if child died, when.
Women who had live birth since last interview, and child in household: due date, baby born early or late, birthweight. GAINS scale basic development question about feeding, sleeping, drying; breastfeeding; if twins, type of twins.

Children aged 3, 5 or 8 (asked of ‘responsible adult’)
For all children aged 3 whether child: has long term health condition; a series of questions about learning physical and cognitive skills, behaviour and psychological wellbeing (Vinelands adaptive behaviour scale).
For children aged 5 or 8 Strengths and Difficulties Questionnaire, child happiness at school, child risk aversion behaviours
For children age 3, 5, 8 parental reading with child, regular meal and bed times

For parents of all children aged under 16 – mother’s return to work plans, child care arrangements, parenting module on the way parents and children interact at home, parenting styles, non-resident parent relationships, child maintenance, etc.
What scientific questions can *Understanding Society* address about the experiences of different birth cohorts in the UK?

(See Pelikh, 2019 for a fuller discussion)

ESRC’s specification for this work focused very much on the role of birth cohorts in *Understanding Society*. We hope we have demonstrated above that *Understanding Society* can provide invaluable insights into children’s lives and their development and transition to adulthood without being a birth cohort. Indeed, it has a number of advantages over the cohort design, for example, the Study includes all family members and follows them all over time, and includes all children as they are born or join the Study, which means children are studied in their family context and inter-relations within families can be examined. *Understanding Society* interviews children of all ages and birth years so that effects are captured in multiple historical contexts and for all cohorts, not just set birth years/time periods, increasing the generalisability of the findings for today’s scientific and policy questions. In the main therefore this think-piece focuses on the value of *Understanding Society* for early life research and child development, but in this section we highlight the ways in which *Understanding Society* can contribute to research from a cohort perspective.

*Understanding Society* was originally designed to include approximately 10,000 respondents per decade of birth years at Wave 1, while maintaining the Study’s representation of the general population, so with naturally smaller cohorts at the oldest ages. Current sample sizes for different birth cohorts are shown in Figure 1. As new members join the sample each year, and with the proposed general population and immigrant boost samples (in wave 13-15 bid, under consideration), numbers in specific cohorts will increase (although at the same time attrition will reduce numbers). In addition new birth cohorts are developing on an ongoing basis, so that *Understanding Society* can already shed light on the childhood experiences of children born in the 21st century (see Pelikh, 2019).

**Figure 1** Birth cohort sample sizes in *Understanding Society*, Wave 2

![Birth cohort sample sizes](source: Pelikh, 2019)

A key advantage *Understanding Society* has for cohort analysis is that it asks the same questions – across a range of topics - of people of all ages over time, so that there is a relatively large spread of adults from different birth cohorts with comparable data. The Study also asks retrospective questions so full lifetime histories can be built up, even if
interviews start at different life-stages. These questions have mainly been asked on initial entry into the Study or in periodic large additional event history data collections, which means some groups are missing these data. Nevertheless, it is relatively straightforward to identify cohort differences in a range of experiences and trajectories over time. To illustrate this Table 1 shows the age of starting work across eight birth cohort decades.

Table 1 Median age at entering the labour market* by cohort and gender

<table>
<thead>
<tr>
<th>Cohort</th>
<th>age started first full-time job</th>
</tr>
</thead>
<tbody>
<tr>
<td>1921-1930</td>
<td>15.2 15.3</td>
</tr>
<tr>
<td>1931-1940</td>
<td>15.7 16.0</td>
</tr>
<tr>
<td>1940-1950</td>
<td>16.3 16.2</td>
</tr>
<tr>
<td>1951-1960</td>
<td>17.0 17.0</td>
</tr>
<tr>
<td>1961-1970</td>
<td>17.7 18.1</td>
</tr>
<tr>
<td>1971-1980</td>
<td>19.3 19.6</td>
</tr>
<tr>
<td>1981-1990</td>
<td>19.3 20.2</td>
</tr>
<tr>
<td>1991-1994</td>
<td>19.8 20.9</td>
</tr>
</tbody>
</table>

Note: *first full-time job or self-employment spell, Waves 1-5
Source: Pelikh, 2019

Action 6: The Study team plan to examine and revise our retrospective histories data collection strategy to ensure we have an effective system for collecting missing histories from participants.

In a literature review we were able to identify 32 papers which undertook cohort analyses based on the Study (Pelikh, 2019). These mainly investigated fertility patterns and union formation among different cohorts, including a number having a specific focus on the ethnic minority samples within the Study. Cohort analyses also included investigation of changing patterns of health and wellbeing, health behaviours, political affiliations, wealth and retirement incomes, transitions to employment and other adult roles, residential mobility and inequalities (Pelikh, 2019). In a significant subset of these papers, data on UK cohorts from Understanding Society are included in cross-country cohort analyses, for example on migrant fertility patterns (Kulu and Hannerman, 2019), or partnerships and health (Perilli-Harris, et al., 2019). Most of these papers focus on cross-cohort analyses of adult outcomes and trajectories, but a few are based on the youth data and focus on youth-adult transitions (Schoon and Lyons-Amos, 2016; Pelikh and Kulu, 2018). In a few examples, Understanding Society data are being used alongside the British Birth Cohorts to provide evidence on the experiences of more recent cohorts, e.g., on social mobility (Bukodi et al, 2015). One different approach to cohort analyses that is only possible within a panel study is to create cohorts based on key events, such as when someone’s first job occurred or they established their first union (Berrington and McGowan, 2014). This provides the opportunities to track cohorts of people over time with different kinds of baselines, providing a different perspective on subsequent life course patterns than are typically possible with birth cohorts per se.
What are the new research opportunities and scope for expanding *Understanding Society* data on early life?

As noted above the basic design of *Understanding Society*, as a household panel study, creates a wide range of opportunities to research childhood in a household and wider family context from pre-conception until old age. However, focused data collection, and use of the Study, in early life has been missing to date.

Much is already known about the impact of early life circumstances on childhood development and later adult outcomes (e.g., see Almond *et al.*, 2018 for a recent review of the large relevant economics literature). However, changing family, social, political, economic and physical environments bring new challenges to growing up in the UK in the 21st century, and limitations in past data mean we have better understandings for some specific life-stages and exposures than others. For example, Almond *et al.* (2018) argue that much more is known about mothers’ role in childhood than fathers’, and very little about multi-generational effects; that research often relies on proxies for the exposures of interest – such as low birth weight for poor in utero development - rather than direct social and biological measures of concerns; and that there is a lack of real-time data collected continuously through childhood to examine critical or sensitive periods in childhood. Finally, they argue that little is known about some aspects of child development such as non-cognitive skills, and new environmental exposures are emerging that have not yet been investigated.

Collecting continuous data throughout life on people’s everyday lives within a family context, enabling inter and intra-generational/family relationships to be explored, is a fundamental feature of *Understanding Society*. To date, our data on children themselves, however, has been relatively modest. Enhancing and expanding data collection in *Understanding Society* around pregnancy and the early lives of children would enable researchers to investigate a range of new scientific questions that are highly pertinent to current policy concerns. These could include:

- How do men and women’s family expectations and fertility intentions change with their changing circumstances over time? How do difficulties with conception influence behaviours and plans?
- How do biology and social factors interact for men and women in different environments and contexts through pregnancy and early life to influence early life outcomes for their child(ren)?
- What are mother and father-to-be’s expectations during pregnancy about parenthood, their roles in their child’s life and maintaining other roles (paid employment, household, social, etc)? How do these change over time, and do they match reality for whom, in what circumstances? How do maternity and paternity policies and the leave taken by both parents influence these expectations and behaviours? What role models and support do men and women draw on as they begin their new roles as parents?
- How do siblings interact with each other as they grow up? How does each parent’s behaviours and attitudes differ for each child and how does this influence the child’s outcomes? How do children perceive their parents’ behaviour towards themselves and their siblings, and how does this affect them?
- How do children’s characteristics and behaviours impact on parental outcomes such as labour market participation, health and wellbeing, income and wealth?
• How do other generations (grandparents) and wider family networks (aunts/uncles, cousins) interact with new parents and their children at different stages of child development? How do children view these other adult role models?
• How do families’ everyday circumstances change in different environments and influence child development (health, social, cognitive and non-cognitive skills, attitudes and beliefs) during childhood and into adolescence and adulthood?
• What role (financial, practical, cognitive) do non-resident parents (and their wider family – new partners and children, grandparents) play in children’s lives? How is this influenced by the nature of the separation (eg if parents never lived together to when in a child’s life separation occurs), parental relationships with the child before separation, geographical distance, both parents’ circumstances and attitudes?
• How do children’s lives, and the kinds of influences highlighted above, differ in different contexts - UK countries and neighbourhoods, ethnic group and generation, social and economic backgrounds, family structures, birth cohorts and time periods?

The questions above illustrate the kinds of research that is already or might be possible with different kinds of enhancements to Understanding Society. Before considering the next steps into how to develop data for such research, we examine who has babies in Understanding Society, and how they reflect the general population.

**Women of childbearing age and new mothers in Understanding Society**

(See Fisher, 2019 for a fuller discussion)

There are two broad issues to assess when considering whether Understanding Society could enable researchers to answer the questions above: Does the study have sufficient parents/children in different groups to enable robust analysis? Are the parents and children representative of the contemporaneous general population?

**Sample sizes**

There is no one definition of sample size in the longitudinal setting. A given sample of individuals will deliver analysis samples of different sizes, depending on the research question and how the researcher needs to treat the repeated observations on the same individual. Fisher (2019) identifies three definitions that are relevant for the majority of analysis types:

• “total analysis sample” which counts repeated observations on the same person independently and so the sample is straightforwardly the sum of the number of units per wave (currently waves 1-8). This would be relevant, say, if a researcher wanted to analyse a large sample of new mothers cross-sectionally and was willing to pool new mothers from all waves.
• “total analysis sample” by wave, this would be relevant if a researcher wanted to calculate trends over time, say the trend in poverty of new mothers, or if they wanted to restrict analysis to a particular set of time periods (waves).
• “unique persons available for analysis” over all eight waves, which would be relevant, say, for studying the living standards of new mothers following child-birth, e.g., when calculating the number of waves that a given new mother spent in
poverty and each mother contributes exactly one observation to the analysis sample.

In Waves 1-8 of Understanding Society the total analysis sample is 93,120 women of childbearing age, 18,103 of whom are poor (defined as where the equivalised household income is less than 60 percent of the median household income). When looking at the unique cases, there are still substantial sample sizes with 25,205 women of childbearing age, of whom 9,188 are poor. A group that has significant policy interest is single mothers. In Understanding Society waves 1-8, the total analysis sample of single mothers is 15,699 and for single mothers of newborn babies the total analysis sample is a sizeable 5,810.

How do Understanding Society mothers compare to the general population?

Fisher (2019) compares the characteristics of mothers in Understanding Society with those in the ONS Family Resources Survey (FRS). He concludes that the two surveys give a remarkably similar picture of the broad characteristics of new mothers, showing that Understanding Society is representative of the UK population of new mums. In relation to living standards at Wave one, Understanding Society mothers look very similar to the FRS ones in terms of age, household size, marital status, ethnicity and region. Figure 3 shows that there is little evidence of difference or divergence in their living standards over time. We conclude that Understanding Society starts out and remains representative and is thus a valuable data source for research on new mothers and babies.
While *Understanding Society* replicates the cross-sectional representativeness of FRS, there are many types of research that can only be done with *Understanding Society*. One major advantage of *Understanding Society* is following people over time. In relation to pregnancy and early child development, a long spell in poverty is likely to be worse than a short one. For example, during short spells of low income, a family might draw down savings or borrow from relatives to fund consumption. This is more difficult during longer poverty spells. Therefore, the cumulative effect on well-being of spending multiple years in poverty can be assumed to be greater than for a period of short duration. Table 2 shows the number of women with newborn babies and women of child-bearing age who have spent time in poverty.
poverty during the first 8 waves of Understanding Society. It shows that over half of women with newborn babies have spent at least one out of eight years in poverty, and 18 percent of mothers with newborns have spent at least half of the 8 years in poverty. Women with babies, on average are more likely to live in poverty and to spend longer living in poverty than all women of child bearing age. There is no external source of similar longitudinal data with which to compare our findings.

Table 2: Number of years spent in poverty

<table>
<thead>
<tr>
<th>Number of years in poverty</th>
<th>Percent with newborns</th>
<th>Childbearing age</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>48.88</td>
<td>55.14</td>
</tr>
<tr>
<td>1</td>
<td>16.58</td>
<td>15.23</td>
</tr>
<tr>
<td>2</td>
<td>9.57</td>
<td>9.5</td>
</tr>
<tr>
<td>3</td>
<td>6.69</td>
<td>5.95</td>
</tr>
<tr>
<td>4</td>
<td>5.97</td>
<td>4.82</td>
</tr>
<tr>
<td>5</td>
<td>4.87</td>
<td>3.33</td>
</tr>
<tr>
<td>6</td>
<td>3.21</td>
<td>2.97</td>
</tr>
<tr>
<td>7</td>
<td>2.62</td>
<td>1.93</td>
</tr>
<tr>
<td>8</td>
<td>1.62</td>
<td>1.12</td>
</tr>
</tbody>
</table>

Number of individuals: 1340 4151

Notes: Sample of individuals who were in the corresponding state (with newborns or of childbearing age) in at least one year.

Source: Fisher 2019

New babies in Understanding Society

Newborn babies are identified in Understanding Society by the initial household interviewee completing information on everyone in the household, and the fertility questions asked of men and of women in annual events, although some of the pregnancies identified in this way may not result in live births or the babies, when born, may not reside in the household. Currently, we include in the sample all new babies resident in the household and for whom we have been provided with a name and date of birth. Some babies born to sample members may live outside the household, other babies identified in the household interview have no further information gathered on them. Currently mothers provide the main source of information on the first year of the baby’s life, and these data are only asked if a woman has had a baby in the last 12 months. There is a slightly higher rate of refusals and proxy interviews among women in the year a baby is born. Table 3 shows the number of new babies identified in each wave and the number of these enumerated at the subsequent wave.
<table>
<thead>
<tr>
<th>Wave</th>
<th>Number of newborns in wave</th>
<th>Of these how many were enumerated at subsequent wave</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wave 2</td>
<td>1118</td>
<td>933</td>
</tr>
<tr>
<td>Wave 3</td>
<td>995</td>
<td>815</td>
</tr>
<tr>
<td>Wave 4</td>
<td>862</td>
<td>745</td>
</tr>
<tr>
<td>Wave 5</td>
<td>733</td>
<td>606</td>
</tr>
<tr>
<td>Wave 6*</td>
<td>553</td>
<td>473</td>
</tr>
<tr>
<td>Wave 7</td>
<td>683</td>
<td>585</td>
</tr>
</tbody>
</table>

*The annual events module was not asked of new sample members joining Study as part of the Immigrant and Ethnic Boost sample

Understanding Society’s sample was designed to capture 1000 births per year at Wave 1; this number then declines with the attrition among women of childbearing age. The proposed general and immigrant population boost samples (alternating every five years) in the wave 13-15 bid (currently under consideration) will regularly increase the sample size of new babies born within the Study. In any future boost we need to ensure that the fertility module is asked in their first wave to identify newborns in the boost. The analysis for this paper has revealed that we are not identifying newborn babies, and collecting key data on them, as well as we could be, and we are investigating this further to identify ways of improving our data capture on this key group. Our proposals around event triggered data collection outlined below will significantly enhance our data at this lifestage. Newborn babies are also lost to follow up if born to women who are temporary sample members who then leave the Understanding Society household (Fumagalli, 2019), a policy we are reviewing.

Action 7: The Study team will maximise the number of Understanding Society babies that are included in the sample by improving our identification of newborns, developing approaches to ensure we collect timely pregnancy and early life data on as full a sample as possible, for example with shorter questionnaires, asking questions about the first year of life to fathers as well as mothers, changing the proxy interview to include these questions, and by collecting data retrospectively in missed in the first years of life. In addition we will reconsider our following rules for newborns born outside the household and/or to TSMs.

Next Steps: Developing the Understanding Society pregnancy, early life and family data

As outlined above, a number of steps are already in train or planned to improve Understanding Society data at this key life-stage. We are reviewing our current questions and data capture to ensure we have the best questions for this life-stage, and that we gain the information, where we can do this effectively, for the maximum number of children. We are better linking our data across households so that multi-generational relationships can more easily be studied, and we are investigating the best way to expand the study to include non-resident parents/separated families. In the Longitudinal Studies Review, there are three further areas identified as crucial to research on childhood where we feel Understanding Society could contribute better data and hence new research opportunities: the more regular inclusion of biological data, event-triggered data collection and more intensive data collection around pregnancy.
Regular biological data collection

As the Longitudinal Studies Review (Davis-Keen, 2017) notes a success of the UK longitudinal studies is their interdisciplinary and innovative approach, in particular including the collection of biomarkers and the study of genetic, epigenetic and other omics. To date Understanding Society has carried out one collection of biological data, which has been extensively used for a wide range of research ranging from purely genetics and biological to mainstream social science and truly interdisciplinary science. To maximise the value of collecting biological data in a longitudinal study, it is vital that repeat measures are collected so that researchers can utilise them to examine the interplay of people’s biological and social factors over time and at different life-stages. We are currently running a large-scale experiment in the Innovation Panel (IP12) on how we can enable participants and social interviewers to collect biological samples themselves. We proposed in our recent bid for waves 13-15 that we roll out the successful approaches within these experiments (once evaluated) on the main study sample, beginning in wave 15 and then repeated every 2 or 3 waves, depending on funding. This will allow researchers to investigate how people’s health changes over time. Specifically, in the content of this think piece, in IP12 we have included the collection of biological measures from children for the first time to measure their stress levels through hormones. Additionally, the regular collection of biomarker data in adulthood would add to our understanding of parents’ health before conception and during their children’s lives and how parent and children’s health impact on each other.

Action 8: The Study team will maximise learning from IP12 so that the inclusion of biomarkers can be rolled out to adults and children in Understanding Society going forward. Once processes for effectively measuring the IP biomarkers are established, we will investigate broadening the accurate participant-led measurement of biomarkers to other measures, and expanding the range of measures we collect in childhood.

Event triggered data collection

A full report on event triggered data collection has been developed concurrently with this think-piece (Jäckle et al 2019). In brief our aspiration is to gather data on key events in people’s lives from them in a more timely fashion in order to collect different and/or more accurate information than is possible in the annual multipurpose questionnaire. Over the last six months we have undertaken a qualitative study of participants’ willingness to engage in frequent short data collection, how frequent to ask the initial ‘event identification’ question and how best to ask the ‘event identification’ question. Building on the lessons from these early experiments, a programme of research to develop effective protocols for the IP and then the main study are being developed. Funding was included in the most recent bid for waves 13-15 to roll out the successful approach at the end of 2021/early 2022.

Specifically in relation to pregnancy and early life, we propose using the event triggered data collection to identify pregnant women, and their partners, so that more time sensitive data might be collected. At whatever stage we identify a woman is pregnant – through event triggered data collection or through the main study questionnaire – we plan to develop a module to ask both parents-to-be about their intentions in terms of their role in the baby’s life and expectations for family life, as well as their plan for paid employment and other roles following the birth. Researchers will then be able to compare these intentions and expectations with previous (pre-pregnancy) aspirations and how these roles change after the
baby is born. In addition we will ask more detailed questions about both parent’s psychological wellbeing at different stages of pregnancy and health behaviours, currently asked retrospectively in the main questionnaire.

Action 9: The Study team will continue to experiment with protocols for event triggered data collection so that the most effective approach can be rolled out (if funded) in 2021. To develop appropriate modules to be asked either in the main study or through event triggered data collection about parents-to-be’s expectations, psychological wellbeing and behaviours during pregnancy and in the early months of the baby’s life.

Data collection in pregnancy and the first year of life

As noted above there are a wide range of ways we can and will improve data in Understanding Society around pregnancy and early life, and ensure we maximise understanding of the family context. However, to create a truly interdisciplinary resource around pregnancy and early life, collecting biological measures on the parents and child - as well as social survey data outlined above - is essential. We have therefore proposed in our Wave 13-15 bid additional face-to-face data collection to collect biological samples from the mother-to-be in pregnancy, some samples to be gathered during labour, and one additional data collection in the first year of life. As a first step in developing our protocol for this enhancement we commissioned Kantar Public to carry out qualitative focus groups with new parents, pregnant women and midwives, to assess the willingness to participate in such a study, and the practical considerations we would need to address. Kantar Public held an online forum with 11 members of the Understanding Society pilot sample who had a child under two years and four focus groups with new mums, dads and pregnant women who were not part of the Understanding Society, as well as one focus group with midwives. The full report of this qualitative work can be found in a companion report to this (Ogunshakin et al, 2019). In hindsight, including participants who were unfamiliar with the Study was not as informative as we might have hoped for. The dominant messages from the research were that detailed and comprehensive information was required by the participants about why biological measures were being included in a social survey, how samples would be used and stored, the ongoing safety of the child’s data and samples, and the need for financial incentives, all of which are standard practice in Understanding Society data collection. While existing Understanding Society participants felt more comfortable with the structure of information and consent sheets, the unfamiliar participants found the depth of detail rather daunting. This is a well taken reminder of the importance of communicating effectively information about the purpose, confidentiality and process of our data collection in multiple ways to ensure all of our participants can engage effectively with it to make informed decisions about participating in different aspects of the Study.

There were some very helpful findings about collecting data during pregnancy and in the early stages of the baby’s life. Parents, and especially expectant parents, were very concerned about how the Study would be sensitive to any pregnancy complications that might occur. Indeed some expressed reluctance to take part in case discussing the pregnancy so openly might actually ‘jinx’ the outcome. There was also concern about any data collection being carried out too close to the due date, given the competing pressures on parents-to-be’s time at that stage. Given this, it seems appropriate to aim data collection during pregnancy for around 6 months, i.e., once the abnormality scans have been conducted but before the due date becomes too imminent. Leading up to delivery, we
anticipated mums-to-be being asked consent for cord blood to be collected during delivery by midwives. The midwives we interviewed felt, with appropriate informed consent, and assuming no complications, this would be an easy request to fulfil since they are taking samples anyway. Another source of data we proposed collecting was data and measures from women’s ‘pregnancy notes’ and other data from regular check-ups during the pregnancy and the ‘red book’ following the birth. Again, midwives felt this would be straightforward, but parents were concerned about highly confidential data that may be part of the record they would not want to be shared (e.g., about abusive relationships). Midwives assured us that such data are not part of the records, so we need to identify effective ways of reassuring parents about the kinds of data that are covered in these records. They were also concerned about the highly sensitive nature of some of the biological samples we had proposed, in particular, sharing hair samples seemed to be of considerable concern. Developing a better understanding of why parents were so wary of one of our more non-invasive collection techniques is required, so that better designs may be developed, for example asking parents to send samples when they first cut the baby’s hair rather than doing so specially for the Study. Overall, the key lesson from this research suggests that we need to work hard to identify minimally invasive ways of collecting data around this ‘challenging and tiring’ time for expectant and then new parents.

**Action 10:** Drawing on key lessons from this first pregnancy feasibility study, the Study team will design more effective materials to explain the value of data collection at this life-stage to parents, and develop protocols for data collection that can be trialled that minimise burden and invasiveness while maximising the key data that can be collected.
References


Booker, C. (Under Review) ' How do early life trajectories impact early and mid-adulthood health outcomes?' Advances in Life Course Research


The Children's Society and Barnado's (2018) Factors affecting children’s mental health over time London: The Children's Society and Barnado's


