

ESRC Longitudinal Studies Review 2017 Workshop summary report January 2017

Background	2
Consultation	2
Workshop	2
Structure	2
Thematic priority areas.....	3
Methodological, technological and cross-cutting issues	3
Workshop outputs.....	3
Day 1 – thematic priority areas	3
Day 2 – methodological, technological and cross-cutting issues.....	4
Next steps	4
Annex 1: Longitudinal studies review workshop participants list.....	5
Annex 2: Longitudinal Studies Review 2017: Workshop programme for participants.....	7
Annex 3: Summary notes from the Longitudinal Studies Review Workshop 9 January 2017.	9
Introduction	9
Discussion group summaries.....	9
Group A: long-term effects of childhood and adult experience.....	9
Group B: demographic shifts, mobilities (social and physical) and ageing.....	9
Group C: health and wellbeing	10
Group D: equality and inequality.....	10
Group E: biosocial research and genomics	11
Group F: diversity and identity.....	11
Annex 4: Summary notes from the Longitudinal Studies Review Workshop 10 January 2017	
.....	12
Introduction	12
Discussion group summaries.....	12
Group 1: design, implementation, participation, engagement.....	12
Group 2: sampling, representativeness, weighting, imputation.....	12
Group 3: new forms of data collection, ethics, biomarkers, wearables, other data.....	13
Group 4: data linkage, ethics, feasibility	13
Group 5: capability/training and statistical methods.....	13
Group 6: discoverability, access, user services and communication	13

*“What panel data?
Does it change? Does it matter?
Can you measure it?”*
 (Haiku from a workshop participant, day 1)

Background

ESRC is currently undertaking a strategic review of longitudinal studies led by an international panel. The review seeks to consider the continued and future scientific and policy-relevant research needs for longitudinal resources, and how these needs could best be met. The review will essentially look at the 'what' and 'how' questions in the area of longitudinal studies, ensuring that future investment enables the scientific community to address critical issues in the most appropriate ways. For more information, see the review specification (<http://www.esrc.ac.uk/news-events-and-publications/publications/corporate-publications/longitudinal-studies-review-2017/>).

Consultation

As part of the review process, ESRC has successfully run two consultative elements that have provided the opportunity for input.

First, the online Longitudinal Studies Review Consultation opened on 3 October and closed on 4 November 2016, receiving an impressive 637 completed responses, demonstrating the strong engagement of our longitudinal studies community, broadly conceived. Information pertaining to the consultation, including the initial report on the consultation findings and an infographic, can be found on the ESRC website (<http://www.esrc.ac.uk/news-events-and-publications/publications/corporate-publications/longitudinal-studies-review-2017/>).

Workshop

The second consultative element relates to the Longitudinal Studies Review 2017 Workshop. This stage was informed by the earlier consultation exercise. On 9 and 10 January 2017, ESRC hosted a workshop at Nuffield College, University of Oxford, which sought to explore:

- What are the future scientific and policy needs for data to address the types of research questions for which longitudinal data has typically been used?
- How can these needs be met and what are key challenges to be addressed?

Participants were asked to take a challenge-led approach that looked beyond what we do currently and were encouraged to be critical, ambitious and creative. ESRC invited experts from across a range of academic, policy and professional communities; from around the UK and overseas, in order to reflect a wide range of key perspectives; and members of the Longitudinal Studies Review Panel and Steering Group (comprising representative of ESRC committees, MRC, GO-Science and Wellcome). The workshop participant list can be found at Annex I.

Structure

The workshop started with a series of presentations that sought to contextualise the workshop, communicate the overall aims and objectives, and set out what was expected of participants. As noted, there was a focus upon critical and creative thinking, and presentations explicitly spoke to this and helped to set the context for the breakout group discussions: the main sections of the workshop that provided opportunities for focused

detailed discussion. The two sets of discussions were broadly focused around a) thematic priority areas and b) methodological, technological and cross-cutting issues. Careful consideration was given to ensure breakout groups included individuals who would challenge normative perspectives and provoke discussion that really probed *why* longitudinal data would be needed in the future and the challenges to be addressed. Please see Annex 2 for the workshop programme.

Thematic priority areas

Participants were assigned to thematic area groups and in each they were asked to identify, scope out, and critically challenge the types of priority strategically important research questions that social science and inter-disciplinary data will be needed to address in the future, and to explore the types of evidence and resources needed to address them. The thematic areas were drawn out of the consultation report:

- Long-term effects of childhood and adult experience
- Demographic shifts, mobilities and ageing
- Health and wellbeing including mental health
- Equality and inequality
- Biosocial research and genomics
- Diversity and identity

Methodological, technological and cross-cutting issues

Participants were assigned to groups and asked to scope out and challenge each priority issue. These issues were drawn out of the consultation report:

- Design, implementation, participants engagement
- Sampling: representativeness, weighting, imputation
- New forms of Data collection, ethics, biomarkers, wearables, other data
- Linkage, ethics, feasibility
- Capability/training and statistical methods
- Discoverability, access, user services, communication

Workshop outputs

Day 1 – thematic priority areas

Each group was asked to consider one the above themes, addressing research resource needs in broad terms (ie not only quantitative data, looking beyond ESRC funded resources, and considering interdisciplinary approaches), and reflecting upon the following questions:

- What are the priority research and policy questions of the future?
- **What** is the **unique value** of longitudinal data to these questions? Without longitudinal data what, if anything, would we be unable to know? What could be done to fulfil the potential of longitudinal resources?
- What **contextual factors** impact the utility of longitudinal studies and other data?

Please see Annex 3 for a summary of discussion group outputs.

Day 2 – methodological, technological and cross-cutting issues

Each group was tasked with scoping and challenging the methodological, technological and cross-cutting issues listed above, considering alternative evidence sources and approaches that are not currently supported by ESRC investments, and addressing the questions:

- Starting in 2017 and looking ahead, what would we use instead?
- What would be missing if we did not have longitudinal studies, and how important is this?
- In relation to the remit of your group, what are the priority challenges for longitudinal studies? What potential solutions should be explored?

Please see Annex 4 for a summary of discussion group outputs.

Next steps

The workshop will inform the Review Panel's thinking on the evidence it needs to procure and assess through a series of workstreams in the next stage of the review. The Panel is currently carefully considering the workshop outputs to draw out key themes and identify cross-cutting issues to address.

The Panel will use the evidence gathered in the review to evaluate where and how life-course data from longitudinal studies can make a unique, valuable and effective contribution to addressing key research questions identified, and to establish priorities, broadly speaking, for the provision of these and related sources. The review will continue throughout 2017 and is expected to report to ESRC in February 2018. Please see the ESRC website for more information (<http://www.esrc.ac.uk/news-events-and-publications/publications/corporate-publications/longitudinal-studies-review-2017/>).

Annex I: Longitudinal studies review workshop participants list

Name	Organisation
Sharon Witherspoon	Academy of Social Sciences
Ruth Gilbert	University College London
Kate Pickett	Born in Bradford
Harriet Barnes	British Academy
Alison Park	Cohort & Longitudinal Studies Enhancement Resource (CLOSER)
Mark Birkin	Consumer Data Research Centre: University of Leeds
Jane Falkingham	Centre for Population Change, University of Southampton
Paul March	Department for Work and Pensions UK Government
Clare Baker	Department of Education UK Government
Peter Evans	Department for International Development UK Government
Rachel Conner	Department of Health UK Government
Nick Goldman	European Bioinformatics Institute
Jane Elliott	Economic and Social Research Council
Bridget Taylor	Economic and Social Research Council
Rebecca Fairbairn	Economic and Social Research Council
Fiona Armstrong	Economic and Social Research Council
Frances Burstow	Economic and Social Research Council
James Fenner	Economic and Social Research Council
Joe Ellery	Economic and Social Research Council
Joy Todd	Economic and Social Research Council
Chiaki Beis	Economic and Social Research Council
Jo Goddard	Economic and Social Research Council
Becky Leithall	Economic and Social Research Council
Michele Erat	Engineering and Physical Sciences Research Council
Tom Wells	Government Office of Science UK Government
Emma Gordon	Her Majesty's Treasury UK Government
Paul Clarke	University of Essex
Rob Joyce	Institute of Fiscal Studies
Sam Clemens	IPSOS MORI
Bruce Hayward	KANTAR Public
Louise Arseneault	Kings College London
Tim Hubbard	Kings College London
Corinna Kleinert	Leibniz Institute for Educational Trajectories, Bamberg Germany
Emily Grundy	London School of Economics
Joe McNamara	Medical Research Council
Kirby Swales	NATCEN - independent Social Research institute
Patrick Sturgis	National Centre for Research Methods
Simon Gardner	Natural Environment Research Council

Pam Miller	National Society for the Prevention of Cruelty to Children
Melinda Mills	University of Oxford
Andrew Dilnot	Nuffield College, Oxford
Teresa Williams	Nuffield Foundation
Matthew Steel	Office of National Statistics
Leon Feinstein	Children's Commissioner
Caroline Kenny	Parliamentary Office of Science and Technology
Anna Vignoles	University of Cambridge
Ioannis Petropoulos	Wellcome Trust
Mary de Silva	Wellcome Trust
George Ploubidis	Centre for Longitudinal Studies
Paul Bradshaw	ScotCen Growing up in Scotland
Nyovani Madise	University of Southampton
Ellen Flint	Department of Work and Pensions UK Government
Gordon Harold	University of Sussex
Andrew Chesher	University College London
Graham Hart	University College London
Matthew Woollard	UK Data Service (UKDS)
Prerna Banati	Unicef Innocenti
Leslie Davidson	University of Columbia, USA
Chris Dibben	University of Edinburgh
Nick Bailey	University of Glasgow
Pam Davis Kean	University of Michigan, USA
Sandra Tang	University of Michigan, USA
Qiang Ren	University of Peking, China
Micheline Beaulieu	University of Sheffield
Louise Ryan	University of Sheffield
David de Roure	University of Oxford
Ray Chambers	University of Wollongong, Australia
Rich Horton	Zentano Group

Annex 2: Longitudinal Studies Review 2017: Workshop programme for participants

9-10 January 2017 - Nuffield College Oxford, OX1 INF

9 Jan	
1200-1300	Arrival and registration – Nuffield College
1300-1345	Lunch in Junior Common Room
1400-1500	Plenary in Lecture Theatre Introduction Andrew Dilnot (Warden of Nuffield): welcome and key note Rich Horton (Facilitator): Introduction to workshop and discussion session 1 Challenges: Emma Gordon (Government Economic and Social Research Team), Teresa Williams (Nuffield Foundation)
1500-1730 with tea break JCR	Group discussions: Thematic areas -in break-out rooms
1730-1800	Plenary in Lecture Theatre Summary reports from Group Chairs, discussion session 1 Summing up on thematic issues from Day 1: Pam Davis-Kean, Panel Chair
1900-1930	Reception and address by Jane Elliott (ESRC Chief Executive) in Senior Common Room
1930-2100	Dinner in Hall
	Overnight accommodation in Nuffield College / Rewley House / Royal Oxford Hotel
10 Jan	
0900-0930	Plenary in Lecture Theatre Introduction to session 2: Rich Horton - Facilitator Challenges: Dave de Roure (University of Oxford), Tim Hubbard (King's College London)
0930-1200 with coffee break JCR	Group discussions: Methodological/technological/Cross-cutting issues in break-out rooms
1200 – 12.45	Plenary in Lecture Theatre Summary reports from Group Chairs, discussion session 2 Summing up on methodological & technological issues from Day 2: Ray Chambers, Review Panel member Next steps
1300-1400	Lunch in Junior Common Room
1400	Departure

Annex 3: Summary notes from the Longitudinal Studies Review Workshop 9 January 2017

Introduction

The following summaries capture the key points from the six discussion groups from the perspective of the ESRC Office. The notes were completed by six different individuals, so there may be some variation in style. The below summaries relate to the following questions:

- What are the priority research and policy questions of the future?
- **What** is the **unique value** of longitudinal data to these questions? Without longitudinal data what, if anything, would we be unable to know? What could be done to fulfil the potential of longitudinal resources?
- What **contextual factors** impact the utility of longitudinal studies and other data?

Discussion group summaries

Group A: long-term effects of childhood and adult experience

- Education and skills (ie impact of home learning environment)
- Inequality/diversity (ie inclusion of population not represented in survey)
- Migration (but this is under different group's remit)
- Pregnancy/early data (ie prenatal exposure and early life data)
- Social/family relation and networks
- Socio/emotional wellbeing (general wellbeing), including critical pathway for intervention etc.
- Technology (ie online/offline and identity)

Other points:

- Using different approaches such as microsimulation
- Involving field agencies early enough (to design study better) also inclusion of intervention

Group B: demographic shifts, mobilities (social and physical) and ageing

- Digitisation, AI, Robotics, Human/ Machine interaction
 - Impact on fertility
 - How the structure of the labour market is/will change (hourglass economy)
 - How this changes relationships/ isolation and impact on cognitive function
- Inequalities and social mobility- we need to look beyond population average to capture views of different groups (eg the Brexit shock):
 - Gender pay inequalities/ labour market/ parental leave issues and policy evidence
 - Differences in economic potential as a result of investment in education, eg impact of having to personally pay university fees and maintenance costs (impact on careers, mortgages, fertility) and what are the costs of these changes on society?
 - Inter and Intra generational mobility (what can government DO about this?)
- Ageing, quality of life:

- Life course at the oldest-old end. How do attitudes and beliefs change?
- At the oldest-old end, how is social interaction changing for those not electronically networked? What is the impact of this (isolation/ cognitive function)?
- Are we becoming our parents? Or do we take our generation's technology with us?
- Will added years be quality years? When does longer life = longer working life?
- What experiences build/ diminish resilience? (interest in different populations)
- Physical mobility:
 - What will happen to the UK population post-Brexit? What will the impacts be?
 - How do physical and social mobility interact?
 - Migration generally- need to know much more about migrants' experience/ resilience and difference between the general population.
 - How do we get the highly skilled out of London?

Group C: health and wellbeing

Three main priority areas:

- Increasing the number of people that age healthily
- Issues of obesity-people are becoming obese earlier and therefore staying obese longer
- Increasing the number of young people with good mental health including research on neurocognitive development

Within these we need to be able to:

- Understand the social, environmental and economic context, links between physical and mental health and the persistence and desistence of problems/factors over the life-course.
- How do different behaviours at different stages in life have different outcomes?
- Identifying points of intervention, tipping points, trajectories etc.
- How to unravel age, period and cohort effects.

Group D: equality and inequality

- What are the drivers of labour productivity in the long term – how does this change as people move between and within educational and working lives? Where might government intervene and how do we use evidence to inform the policies around these issues?
- How does the legacy of violence and conflict affect attitudes and life chances? This is a wide question, ranging from war on an international platform, to localised violence and conflict at a household level. Can these be linked and can similarities be drawn between them, providing international learning in ways of assessing and dealing with them?
- How are our life chances shaped by early life experiences, such as parental investments (in time, money) and does intergenerational transfer affect society – ie: what is handed down to us by older family members, versus what is achieved or attained in our own capacity?
- Hidden populations; those at the extreme ends of society such as: offenders, homeless, drug abusers, immigrants. These are often the first to drop out of a

longitudinal study, or are not included at all, and are poorly understood. How can we begin to understand these issues better and gain insight into the cascade of events that may lead to these types of situations?

- How is technology changing the world of work and with what effects on the distribution of living standards? What are the issues in inequality in the digital era - access, training, and understanding? What is the role of social media in increasing the inequality gap in our lives?

Group E: biosocial research and genomics

- Interaction of social exposures with the genome and physiology, including abuse/neglect.
- Inform interventions: who the interventions work for and intervention evaluation.
- Effects of toxic stress on different groups.
- Trajectories for immigrant populations and how they compare to non-immigrants.
- Health and other outcomes for immigrant populations and minority groups.
- Possible policy questions around the genetic effects of people in prisons – which may inform policy of which prison environment to place people in.
- Which are the high risk populations and why are some individuals more sensitive to exposures and stressors? Why are there different outcomes for different people?
- Policymakers want evidence on what is happening now (live/real-time data could provide this if it was accessible and quality-controlled).
- Genetics and genomics are relevant to all questions.
- Should genetics inform policy? Are there risks to the targeting of interventions and not targeting those who do not express adverse outcomes?
- Validation and enrichment of big data / social media data.

Group F: diversity and identity

- Social relationships (longitudinal data are uniquely positioned to capture the changing nature of social relationships):
 - Why do people make certain choices?
 - How do family relationships change and influence individuals (ie interrelationships in childhood)?
- Migration, identity and cohesion: migrants are difficult to capture due to their mobile nature and longitudinal data can be uniquely placed to follow migrants, if designed with this in mind, for example over periods of 5-10 years as opposed to over the whole life course (appreciating limitations that relate to mobility/movement).
- Data to inform services: this relates to scientific questions but also has a research design element. The needs of services (ie what needs can the data address) and the potential uses should be defined at the outset so data is directly applicable to service design.

Annex 4: Summary notes from the Longitudinal Studies Review Workshop 10 January 2017

Introduction

The following summaries capture the key points from the six discussion groups from the perspective of the ESRC Office. The notes were completed by six different individuals, so there may be some variation in style. The below summaries relate to the following questions, framed by the statement ‘*imagine we had no longitudinal studies*’:

- Starting in 2017 and looking ahead, what would we use instead?
- What would be missing if we did not have longitudinal studies, and how important is this?
- In relation to the remit of your group, what are the priority challenges for longitudinal studies? What potential solutions should be explored?

Discussion group summaries

Group 1: design, implementation, participation, engagement

Development of longitudinal data spine was a recurrent theme in the group. Broadly two options were identified: preferably, a spine made up of linked administrative data including health data, that would be supplemented by survey data including (but not limited to) longitudinal surveys which could be of particular sub-populations or for particular purposes (such as sensitive or qualitative information), OR a spine made up of longitudinal survey data (potentially including non-mainstream LSs) supplemented by linked administrative data. Looking at the potential use of the Census Longitudinal Studies was also suggested. The realistic feasibility of data linkage in the near future is therefore a critical question. A range of types of studies could be connected with the spine. The possibility of a new birth cohort on the conventional model was raised and the feasibility questioned, but this was not pursued in detail.

Priority challenges for longitudinal studies were considered as:

- Mental health research questions
- Capacity for data collection and use including skills and methods. The idea of a longitudinal studies foundation was thought to have potential value in addressing capacity and awareness.
- Impact of childhood on later outcomes; family structures, context, dynamics; migration and trajectories

Group 2: sampling, representativeness, weighting, imputation

The most prevalent discussion topic was about representativeness, of which the group decided it had a relative and absolute meaning, and could also be viewed in an objective vs. subjective manner. The group suggested that we would develop a new longitudinal study if these did not exist (perhaps utilising administrative data), however, noting that in order to make data comparable across time fixes need to be developed to ensure that new and older data could be compared: in order to futureproof longitudinal data. The idea of the ‘Big Dynamic’ was suggested to rethink the way we structure and use longitudinal studies.

Group 3: new forms of data collection, ethics, biomarkers, wearables, other data

The group agreed that there should be a 'backbone' of longitudinal data to link other data to, but did not agree whether the longitudinal studies provide the best 'backbone'. The suite of national longitudinal studies provides a backbone with which to integrate other data with a contextual approach, which would be lost without the studies, although there may be an alternative 'backbone'. Rigour of measurement and prospective studies would be lost without the longitudinal studies.

We need an ability to harness new knowledge and technologies in longitudinal data to identify early indicators and enrich existing datasets. There are uncertainties around the value of new forms of data in studying individual change and experience. The volume and richness of new data becoming available brings both challenges and solutions. There needs to be a readiness to work with new forms of data and agreed approaches to the ethics and methods for using legacy or footprint data (eg data on deceased people).

Group 4: data linkage, ethics, feasibility

Funding Google and/or other big data companies was seen as a potential option, however, issues of access and consent were discussed, as well as a question surrounding whether this data could enable the answering of social science questions. Admin data was seen as a feasible option (with Scandinavian countries mentioned as the gold standard) which could be utilised to establish 'electronic cohorts' and linked to commercial data, whilst acknowledging that gaps would exist (eg household composition, community cohesion, attitudes and aspirations). More reactive and in-depth longitudinal studies (narrower) were suggested, with admin/commercial data filling in gaps, however, we would lose out on the unforeseen benefits of longitudinal studies data (ie things we don't know we will want to know about – that are captured by the depth of longitudinal data). We might also need to consider issues with the representativeness, quality and coverage of admin and commercial data.

Group 5: capability/training and statistical methods

This group recognised that the data collected in other ways is already far bigger than we can ever collect through Longitudinal Studies, but this data lacks the representativeness and context needed to make it meaningful. Linkage is the way forward. Once we can properly link data, non-longitudinal studies data would allow faster information access (including on micro-changes) while longitudinal studies would provide a contextual basis for that data. Linkage is, however, a long way from reality and in the meantime, longitudinal studies are an expedient approach to accessing data we can't (for practical, system and legislative reasons) access in other ways.

The group also reflected on the apparent lack of entrepreneurialism in our approach to longitudinal studies. Why can't we create something that companies or government or respondents come to us to access? Companies are able to get people to give them their data - there are lessons around engagement to prevent attrition for social science.

Group 6: discoverability, access, user services and communication

This discussion was largely centred on administrative data and linkage and the Swedish model was again held in high esteem, however, it was acknowledged that there is a different culture in Scandinavian countries compared to the UK. Whilst there are limitations of using administrative data it was seen as a major alternative. In particular, admin data was seen as lacking contextual information (could richer questions be asked when collecting administrative data in order to address this?). In reference to admin data, the census was

proposed as a 'core' dataset to link additional data to, individual identifiers would enable more accurate linkage, supplementary tailored studies would enable particular transitions to be captured and could provide useful contextual information (longitudinal and cross-sectional). It was noted that there's still a very risk averse attitude towards linkage. The group noted that longitudinal studies provide rich contextual information, enable causality to be captured, have strongly engaged/involved participants, are comparable (particularly across longitudinal studies – although more work needs to be done to harmonise), and capturing information that would fall outside of administrative datasets (ie if individuals do not come into contact with the welfare state they would not be recorded in admin data – but equally, admin data is better at capturing certain groups, such as offenders, homeless people and drug users).

Priority areas were considered to be: data access; data movement and security; understanding the use of longitudinal datasets in order to understand their value; training researchers to access data (and educating why these data sources are important).