

# Growing Up in England dataset

## *Key messages from the ADR England Children & Young People Representative Panel*

February 2022

## Abstract

This report records the first meeting between the [Growing Up in England \(GUIE\)](#) project partners and the [ADR England Children & Young People Representative Panel](#) to discuss the potential of the GUIE dataset. It captures the main findings of the discussion with the intention to inform the future use of the GUIE dataset, including research priorities, scope for growth of the dataset, how to facilitate the voices of children and young people in shaping the direction of research, and more. Stakeholders highlighted more research into children's family background, underrepresented communities and the impact of children's local networks as key areas of research focus. This meeting was held on Monday 4 October 2021; a full list of attendees and areas for future research can be found in Appendices 1 and 2, respectively.

# 1. Introduction

Engagement with stakeholders, such as policymakers, service providers and public advocacy groups is essential to ensure that research using public sector data is truly in the interests of those it aims to benefit. Each stakeholder has an interest in knowing whether the focus and methods of research using public sector data are ethical, robust and useful, and that any potential negative consequences have been considered and mitigated.

On 4 October 2021, a range of stakeholders came together to discuss the Growing Up in England (GUIE) dataset, a new linked data resource (see Box 1) for research in the public interest. The purpose of the event was to provide an opportunity for stakeholders to ask questions and share their views on the GUIE dataset, and to suggest priority areas for its use in future research for the public benefit.

The event consisted of an introduction of the GUIE dataset by representatives from the Office for National Statistics (ONS) and the Department for Education (DfE) and a facilitated discussion with the ADR England Children & Young People Representative Panel (hereafter referred to as 'the Panel'). The Panel consists of members with the knowledge and expertise to represent the interests of children and young people in discussion around their data and its research uses. A full list of attendees is given in Appendix 1. This report covers the key messages arising from discussions at the event.

## ***Box 1: Overview of the GUIE dataset<sup>1</sup>***

The GUIE dataset enables a more comprehensive understanding of how factors such as family background, school type and geography shape outcomes. Around seven million records have been matched, producing a significant sample size for analysis. The GUIE dataset links two existing data sources:

- **2011 Census** – which provides a detailed snapshot of the population, including housing information, individual characteristics and geographical data.
- **Feasibility All Education Dataset for England (AEDE)** – a bespoke extract of DfE data which includes five years of attainment information for a cohort of children who were in Key Stage 4 and Key Stage 5 in the 2001/02 academic year, up to 2014/15 academic year, from local-authority maintained schools in England.

Much of the de-identified GUIE dataset is already available to external researchers via the ONS [Secure Research Service \(SRS\)](#), to facilitate research to improve policymaking for children's health, education and wellbeing across a range of health and social science disciplines. Researchers will need to be accredited and submit a successful application to access the data.

## 2. The research value of the GUIE dataset

Discussion was guided by the potential of the GUIE dataset. Its main objective is to create population-wide data covering roughly 15 academic years, from 2001/2 to 2014/15. Insights derived from this dataset will enable research that aims to understand which factors influence education attainment, including addressing knowledge gaps related to the relationship between income vulnerability and education data.

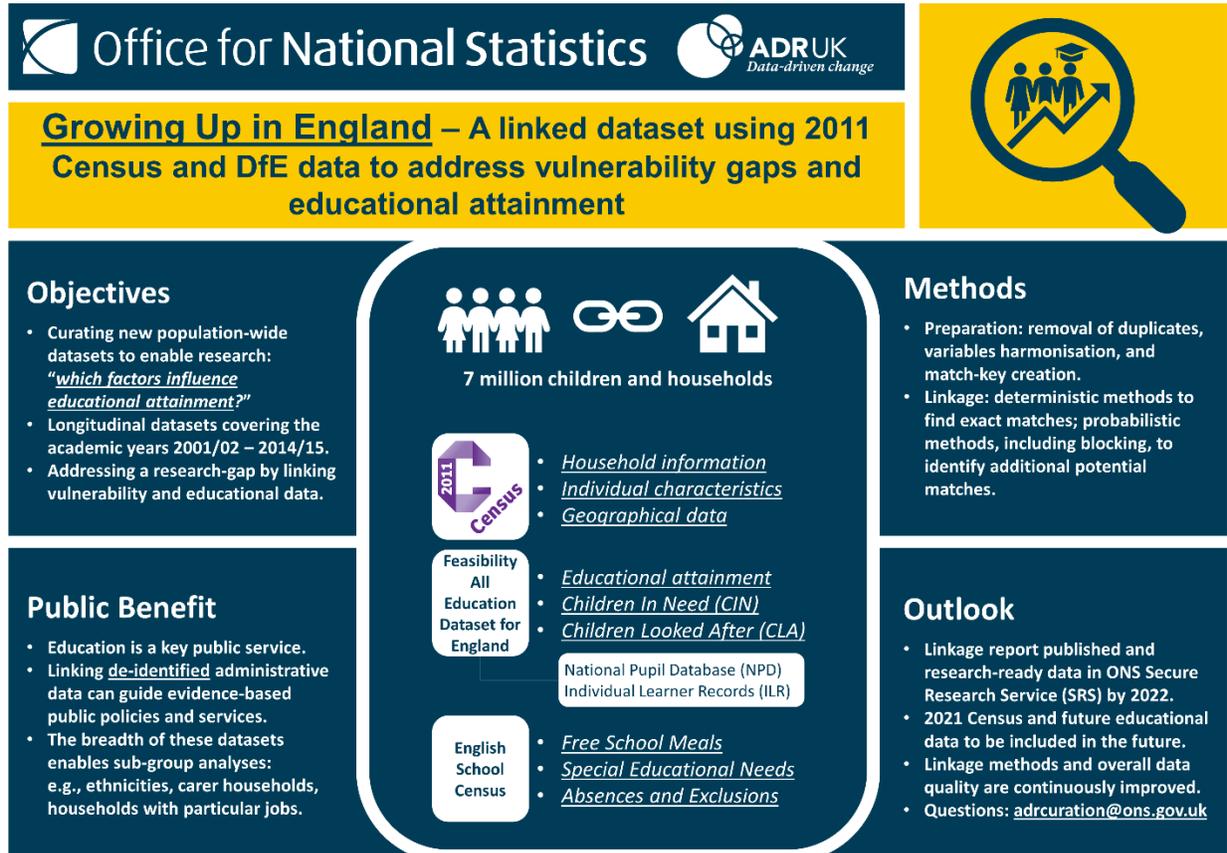


Figure 1: The GUIE dataset

A primary motivation behind creating the GUIE dataset is to provide an evidence base for public policy decision making and improving public services. Within the Feasibility All Education Dataset for England (AEDE) are multiple education datasets, which when linked together, can uniquely fill in knowledge gaps related to education and outcomes. For instance, the National Pupil Dataset (NPD), which forms part of the AEDE brings together various de-identified pupil-level information held within DfE, such as attainment data and data from the school census, enabling data collection at various points in time throughout one’s educational career. NPD data therefore enables analysis of longitudinal outcomes of an individual from early years to higher education. When this data is linked to Census data, for instance, it can offer the potential sub-group analysis, such as on the relationship between education and ethnicity, strengthening its value for policy-relevant research.

Researchers working on the GUIE dataset hope to publish a methodology report next year. Additionally, much of the GUIE dataset is already available to accredited researchers to apply to use in the ONS Secure Research Service (SRS). The 2011 Census, along with future waves of education data, will be consolidated into the linked GUIE dataset ensuring the information remains relevant and up to date. ONS emphasised public benefit is the primary driver of the GUIE dataset and of 'any data existing within the SRS'.

## 3. Research priorities for the GUIE dataset

### 3.1. *Family background*

The Panel raised the importance of the relationship between parents' educational backgrounds and whether a child is likely to continue onto higher education. It was also noted that this might be more prominent for disadvantaged children, such as those on free school meals, for children with disabilities, those who speak English as a second language, and those who receive special needs education. An area of research interest therefore, is whether or how parental background might impact educational opportunities. ONS and DfE clarified that within the GUIE dataset is information on the highest education of parents and their language proficiency, in addition to special characteristics for children, confirming there is scope to explore these areas of interest within the dataset.

It was also noted that some parents seem to experience a lack of confidence when it comes to educationally supporting their children, sometimes manifesting in an inability to relate to their child's educational journey, particularly in relation to higher education. The Panel argued that this issue is rather nuanced, as it deals with both how parents are able to support their children entering higher education, but also how this process can be experienced as a culture change within the family unit, particularly if it is the first child in the family to go to university. The GUIE dataset can be used to quantitatively explore the extent to which qualification levels of parents are associated with children's attainment. However, further exploration into nuances that arise from lived experience, relating for instance to parental confidence in educationally supporting their children, is more suitable for qualitative research.

### 3.2. *Underrepresented communities*

The value of data linkage not only enables us to understand relationships between information from various datasets, but also fills in gaps generated during the data collection process. A question put to the Panel was how we should direct our attention to communities who are often missed or underrepresented in research. While these groups do not make up a majority of children and young people, the need is acute among those in particular circumstances, including children and young people living in temporary accommodation, seeking asylum or with temporary accommodation status. A representative from the Magpie Project noted children experiencing these circumstances aren't necessarily captured well in data from free school meals or pupil premium as they are either not always eligible for these programmes or prefer not to self-report. Understanding the extent to which we can generate insight into the needs of underrepresented communities through this dataset will be important and may point towards further research where there are key gaps.

Concern was also raised about addressing the needs of children and young people in care, particularly if their parents do not fill out the Census, as it can therefore be difficult to factor in household circumstances. A representative from the Access Project observed in their work a demographic of children and young people who are not in the care system or defined as children in need but whose circumstances mean they often moved between family and friends in temporary accommodation. This experience tends to impact significantly on their educational progress but is difficult to track as these children and young people are not always identified by school leaders or included in particular datasets. Meanwhile, a representative from Friends, Families and Travellers noted that the Caravan Count offers more accurate information on the Roma, Gypsy and Traveller

community as opposed to the Census. It was also suggested that linking the GUIE dataset to community services which ramped up during the pandemic could provide a more holistic picture of children and young people's mental health and wellbeing.

### **3.3. Local networks**

Local networks, including communities, local services, and family and support networks, can play an important role in the wellbeing of children and young people. A representative from the Youth Futures Foundation emphasised the importance of looking beyond the immediate household and factoring in the wider community. They argued the Census can offer a narrow nuclear view of a household (up to six people) which isn't relevant to all groups, especially those from an ethnic minority background who are more likely to live in intergenerational households. It was stressed that the Census alone cannot capture the impact of a child or young person living across two households, such as in the case of separate families. Evidence on the extent to which this dataset can help us understand the role of local networks would be valuable and may require further research and additional linkages to fully understand the role local networks play in the outcomes of children and young people.

Meeting with the ADR England Children & Young People Representative Panel reiterated the value of public engagement in the development and use of our linked administrative datasets. The need for the GUIE dataset is clear, as analysis of lone datasets struggle to capture the many factors that determine the outcomes of children and young people. The GUIE dataset will shed light on some of the questions and concerns raised at this meeting, generating vital insights to improve policymaking for children and young people in England. Despite this bespoke and multifaceted resource, there is scope for growth with the addition of new datasets to address knowledge gaps and ensure no one is left out of this crucial programme of work. Key to this will be facilitating the voices of children and young people in shaping the direction of research and publishing accessible (non-academic) outputs. Members of the ADR England Children & Young People Representative Panel have expressed interest in helping ADR UK achieve these aims.

## Acknowledgements

The GUIE project is a collaboration between the [Department for Education](#) and the [Office for National Statistics \(ONS\)](#). It is funded by [ADR UK \(Administrative Data Research UK\)](#). ADR UK is funded by the [Economic and Social Research Council](#) (part of [UK Research and Innovation](#)).

## Authors

Shayda Kashef, ADR UK Strategic Hub, [shayda.kashef@ukri.org](mailto:shayda.kashef@ukri.org)

Gregory Meredith, ADR UK Strategic Hub, [gregory.meredith@ukri.org](mailto:gregory.meredith@ukri.org)

Visit the [ADR UK website](#)



[@ADR\\_UK](#)



## Appendix 1: List of attendees, ordered alphabetically by organization<sup>1</sup>

Chair: [Gregory Meredith](#), Administrative Data Research UK

<b>GUIE project partners</b>	<b>Representatives</b>
Administrative Data Research UK	Christine Boase, Strategic Lead for Communications & Campaigns
Administrative Data Research UK	Gregory Meredith, Senior Research & Impact Manager
Administrative Data Research UK	Bogusia Wojciechowska, Research & Impact Manager
Administrative Data Research UK	Shayda Kashef, Public Engagement Manager
Department for Education	Chris Douglass, Head of Data Sharing and Data Platform
Office for National Statistics	Simon Heckenmueller, ADR Data Curation Research Officer
<b>ADR England Children &amp; Young People Representative Panel</b>	Representatives
<a href="#">The Access Project</a>	Penelope Webb, University Access Officer, West Midlands
<a href="#">The Children's Literacy Charity</a>	Claudette Taylor, Schools Programme Manager
<a href="#">Friends, Families and Travellers</a>	Tommy Buck, Projects & Policy Officer
<a href="#">GenerationR Alliance</a>	Jennifer Preston, PPI & Engagement Priority Lead
<a href="#">The Magpie Project</a>	Jane Williams, Founder & CEO
<a href="#">NSPCC</a>	Susan Cooke, Head of Research & Evidence
<a href="#">Shine Trust</a>	Eleanor Heathcote, Programme Manager
<a href="#">Wellchild</a>	Tara Parker, Director of Programmes
<a href="#">Youth Futures Foundation</a>	Catherine Fitzgerald, Research Manager

---

<sup>1</sup> Invitations were also extended to other members of the ADR England Children & Young People Representative Panel who were unable to attend.

## **Appendix 2: Areas of interest for future research using the GUIE dataset identified by the ADR England Children & Young People Representative Panel**

### **Family background**

- What impact, if any, do parents' educational background have on their children's educational outcomes?
- What support can be offered to families with parents lacking background in higher education if their children wish to enter higher education?
- What factors can positively or negatively impact the educational support children and young people receive from their family unit?
- What is the likelihood a young person from a family with no previous higher education attainment will enter the higher education system? And what factors affect this?

### **Underrepresented communities**

- What data is needed to ensure underrepresented communities are accounted for in the GUIE dataset?
- How can we best support the educational needs of children in care?

### **Local networks**

- How can insights derived from a child's local networks (communities, local services, and family and support networks) impact their educational outcomes?
- What data is needed to best capture children living in intergenerational or multiple households?