Big Data in Education Research: Progress & Prospects

Mike Osborne, Catherine Lido, Phil Mason
Keith Kintrea & Muir Houston

Urban Big Data Centre,
School of Education, Urban Studies
1. Introduction & Background

2. Transforming Learning Cities: Using Glasgow ‘Big-Data’ to challenge notions of ‘Lifewide Literacies’

3. Place-based Drivers of Inequalities: Evidence from Secondary School Attainment in the Greater Glasgow Area

4. Future Research
Educational Disadvantage & Place Team

Good Places – Happy Healthy Citizens

Associations of Lifelong Learning & Place with....

Health
Jobs
Engaged Citizenry
Longevity

Prof Mike Osborne
Prof Keith Kintrea
Dr Catherine Lido
Dr Muir Houston
Dr Phil Mason
Transforming Learning Cities: Integrated Multimedia City Data (iMCD) Project @UBDC

Open Data ‘Product’
1. Survey- 1500 Households
2. GPS Sensors
3. Lifelogging Cameras
4. Social Media Capture

Project Mngr: Dr. Mark Livingston
‘Understanding Glasgow Survey’

Review of National/ EU Surveys + UNESCO Indicators

Attitudes, literacies & Behaviours

- Education/skills
- Sustainability
- Transport
- Cultural/civic
- ICT/technology

Stratified Random Postcode Sampling (Ipsos MORI)

n=2,095, 16-102 years (M= 49.42, SD=19), 45.7% Female, 54.3% Male
Transforming Learning Cities: For engaged older Learners in Glasgow

Regression: Older adults engaged less in ALL learning types

1. **Age** (predicts all learning types)
2. Feeling **Safe** Walking at night
3. Sense of **Belonging** to area
4. **Place**- Local Authority

Transforming Learning Cities: Using Glasgow ‘Big-Data’ to challenge notions of ‘Lifewide Literacies’

Results: Health Literacy mediates effect of Social Support on General Health

Sobel test = 7.84, p < .001
IMPACT: Knowledge Exchange (@Ikea)
Place-based Drivers of Inequalities: British Council school programmes in Scotland, An Impact Study (Nov. 2018)

Livingston, Doherty, Lido, Gale, Parker & Cassar

- Qualitative investigation of n=7 case study schools *(mature & recent, demographically & regionally diverse)*- demonstrating varied engagement histories, motivations, enactments & contributions.
- Quantitative open data analysis of national data
- Quantitative in-depth analysis of Greater Glasgow Area
Quantitative Phase Open Data findings

BC Engagement (2007 – 2017; total #programmes**) is weakly, but significantly correlated with:

• **Deprivation** (SIMD) for most recent year (2016/17):
  Higher engagement in less deprived data zones \( \rho \) (rho)= .10, \( p=.001 \)

• **CfE benchmarks**

• **Effects of BC Engagement remain significant, even when controlling for Deprivation (SIMD)**

<table>
<thead>
<tr>
<th>Effect (Rho/R_s)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>English Reading, Engagement</td>
<td>.22</td>
</tr>
<tr>
<td>English Writing, Engagement</td>
<td>.27</td>
</tr>
<tr>
<td>English Listening &amp; Talking, Engagement</td>
<td>.16</td>
</tr>
<tr>
<td>Numeracy, Engagement</td>
<td>.21</td>
</tr>
</tbody>
</table>
Quantitative Phase Case Study findings

• Case Study of Greater Glasgow: 8 LA secondary schools analysed in eDRIS Safe Haven (attainment challenge & non-LAs compared):
  • n=107 (80.4% w/ some engagement, 19.6% w/ none)
Engagement X Attainment Challenge LA Interaction on SCQF Level 5* & 6** Outcomes

Estimated Marginal Means of Mean number of passes at that SCQF level

Engagers
- Low engagement (1-2)
- High engagement (3+)

School in Attainment Challenge LA

Covariates appearing in the model are evaluated at the following values: SIMD 2012 vigintile: overall rank; school = 9,398
Place-based Drivers of Inequalities: Spatial organization of pupil and school data

- Pupil
- Home neighbourhood
- School
- School neighbourhood
- Placement requests
- Catchment area
Secondary education data

Pupil: attainment S4-S6
  Standard grade general & credit (S4 & S5)
  Higher (S6)

Pupil: attendance & absence

Pupil: leaver destinations

Pupil’s home neighbourhood:
  SIMD, urban-rural indicator, greenspace access

School: staffing

School: main home language

School: Roman Catholic vs. non-denominational

School neighbourhood:
  SIMD, urban-rural indicator, greenspace access

Pupil: personal, demographic & home location

Home-school distance

School: location
### Initial exploratory findings:
4+ passes at SCQF level 6 (Highers) 2014/5

<table>
<thead>
<tr>
<th>Variable</th>
<th>Category</th>
<th>Odds ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
</tr>
<tr>
<td>Gender (male)</td>
<td>Female</td>
<td>1.26</td>
<td>1.16</td>
</tr>
<tr>
<td>Age (&lt;=15 y)</td>
<td>16 y</td>
<td>1.02</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>17 y</td>
<td>0.06</td>
<td>0.04</td>
</tr>
<tr>
<td>SIMD quintile: home (1- most deprived)</td>
<td>2</td>
<td>1.45</td>
<td>1.27</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.84</td>
<td>1.61</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>2.40</td>
<td>2.09</td>
</tr>
<tr>
<td></td>
<td>5 – least deprived</td>
<td>3.99</td>
<td>3.50</td>
</tr>
<tr>
<td>SIMD quintile: school (1- most deprived)</td>
<td>2</td>
<td>1.18</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.52</td>
<td>1.31</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.25</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>5 – least deprived</td>
<td>2.01</td>
<td>1.73</td>
</tr>
<tr>
<td>Urban-Rural: School (large urban area)</td>
<td>Other urban area</td>
<td>0.76</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>Non-urban area</td>
<td>0.67</td>
<td>0.59</td>
</tr>
<tr>
<td>School denomination (non-denominational)</td>
<td>Roman Catholic</td>
<td>0.81</td>
<td>0.74</td>
</tr>
<tr>
<td>Park/garden access &lt;10 min of school (no)</td>
<td>Yes</td>
<td>1.16</td>
<td>1.07</td>
</tr>
</tbody>
</table>
Future Research:
Four strands of Phase 2 research:

1) Place-based Inequalities and Secondary School Attainment and Progression

2) Further Education: Fulfilling Its Purpose?

3) Inclusion in Higher Education: Learner Success in Access, Attainment, and Entering Skilled Employment (with a focus on non-traditional entrants)

4) Inclusive Learning Cities
Can @UrbanBigData help you with your education research?

- iMCD data
- UCAS/ HESA data
- Phone/ Cycling data
- Skills training?
- Future funding ideas?