

**‘Intractable’ schools:  
can an Ofsted judgement  
prevent sustainable improvement?**

*15th August 2019, EARLI Conference*

Dr. Bernardita Muñoz, Dr. Melanie Ehren and Jo Huthinson  
b.munozchereau@ucl.ac.uk

- Ofsted plays an important role in the assessment and functioning of the English school system. Its detailed scrutiny, judgement of school quality and subsequent monitoring of schools' improvement plans is assumed to be 'a force for improvement' (Ofsted 2017, p.5)
- Ofsted uses national test and examination results to classify schools in a four-point scale: Outstanding, Good, Require Improvement and Inadequate. Schools with failing inspection outcomes ("RI" or 'Inadequate') are urged to improve.
- Policy: A small number of schools (124: 72 primary and 52 secondary) have been consistently classified as less than good since 2005. These schools were named 'intractable' by Her Majesty's Chief Inspector (HMCI) (Ofsted, 2017).
- Personal experience: My children attend an inner city state primary school. After it was downgraded from 'Good' to 'RI' by Ofsted in 2015, the most experienced teachers left, enrolment dropped, the head teacher resigned, and the community morale felt generally low. My anecdotal evidence suggests that the RI classification acted as a barrier for improvement.

# Population



| Type of School              | Requires Improvement |           | Inadequate |           | Total |
|-----------------------------|----------------------|-----------|------------|-----------|-------|
|                             | Primary              | Secondary | Primary    | Secondary |       |
| Academy Sponsor Led         | 18                   | 15        | 6          | 14        | 53    |
| Community school            | 13                   | 3         | 13         | 3         | 32    |
| Academy converter           | 7                    | 5         | 2          | 3         | 17    |
| Voluntary Aided School      | 2                    | 2         | 4          | 1         | 9     |
| Foundation School           | 2                    | 4         | 0          | 2         | 8     |
| Voluntary Controlled School | 3                    | 0         | 2          | 0         | 5     |
| <b>Subtotal</b>             | 45                   | 29        | 27         | 23        | 124   |
| Region                      | Primary              | Secondary | Primary    | Secondary |       |
| Yorkshire and the Humber    | 7                    | 7         | 7          | 4         | 25    |
| East Midlands               | 6                    | 4         | 8          | 5         | 23    |
| West Midlands               | 11                   | 1         | 5          | 2         | 19    |
| South East                  | 9                    | 3         | 2          | 1         | 15    |
| North West                  | 3                    | 5         | 0          | 6         | 14    |
| East of England             | 4                    | 2         | 3          | 2         | 11    |
| South West                  | 4                    | 0         | 2          | 2         | 8     |
| London                      | 0                    | 4         | 0          | 1         | 5     |
| North East                  | 1                    | 3         | 0          | 0         | 4     |
| <b>Total</b>                | 45                   | 29        | 27         | 23        | 124   |

- Understand patterns of change over time in this particular group of schools
- What are the characteristics of ‘intractable’ schools?
- Why Ofsted judgements haven’t led to improvement of these schools?
- How is the overall judgement of RI or Inadequate related to judgements of underlying indicators in the Ofsted framework (Leadership and management; Teaching, learning and assessment; Personal development, behaviour and welfare; and Outcomes for pupils)?
- Has the Ofsted judgement contributed to the ‘intractable’ schools’ pattern of lack of change or decline and, if so, how?
- How do school staff, parents and governors of ‘intractable’ schools perceive the validity and fairness of Ofsted inspections, and what are their views on how inspections can support change of their schools?

- Sequential Explanatory Mixed Methods Design (SEMMD) (Creswell and Creswell, 2017)
- Phase 1: quantitative analysis of secondary data sets to capture patterns of change in school performance, student and staff profile, and school context. These patterns will inform our selection of schools for Phase 2
- Phase 2: Multiple-site case study design in 16 schools to explore qualitatively whether and how the classifications of RI and Inadequate have prevented improvement of schools, and how school staff, parents and governors perceive the validity and fairness of their Ofsted inspections

- Ofsted management information records for inspections ('intractable' status and performance of neighbouring schools) 2005-2018
- School performance (value-added progress) and pupil demographics (pupil performance, intake characteristics and volumes) 2005-2018
- School Workforce Census (SWC) Teacher turnover at the school level built from pupil teacher ratio, teacher vacancy numbers and temporarily-filled posts) 2010-2018
- 'Get Information About Schools' (formerly Edubase) for School governance (type, trust and sponsor status), geographical location, area deprivation (IDACI), School financial stability (yearly income and expenditure) 2013-2018

- **Step 1-Propensity score matching (PSM):** Comparison group: sample of schools similar (i.e pupil demographics, value-added progress measures) to the 'intractable' 124 schools in 2005 but that improved their Ofsted grade to understand potential differences between the two groups and how they come to vary over time.
- **Step 2-Cluster analysis:** All schools 2005-2018 that have been less than good since 2005 to understand if there are typical sets of schools in the data, and if 'intractable' schools are overrepresented in one of the clusters and can be distinguished on other data and features than the Ofsted classification, or whether they are a more heterogeneous group, after adding control variables about neighbouring schools' performance.
- **Step 3-Path analysis:** number of years/inspections for which each school has been judged less than good to understand patterns over time. We will construct an ordinal measure of 'intractability' to analyse multiple trajectories of change and test hypotheses about the relationships between and across factors identified as important in steps 1 and 2, and persistence of Ofsted judgements below good.

Multiple-site case study design in 16 schools: 10 intractable (5 primary, 5 secondary) and 6 comparison (3 primary and 3 secondary)

Semi-structured interviews and focus groups with head teachers, teachers, parents, governors and document analysis to explore:

(1) How has the designation of RI or Inadequate affected the school, according to school staff, governors and parents? (

(2) Which improvement plans have been implemented to address the RI or Inadequate classification and how successful were these?

(3) Have schools received any external support and, if so, why do they think it has/hasn't led to improvement?

(4) Which structural changes (e.g. academisation) and regional conditions (e.g. unemployment, low expectations) do schools perceive as contributing to the change/ lack of change?

Analysis: inductive approach based on constant comparisons. Intra-case analysis (within schools) followed by inter-case analysis (between schools and groups). (Miles, Huberman and Saldaña 2014)



- Accountability systems that rank schools based on their effectiveness produce ‘winners’ and ‘losers’ (Riddell 1997; Chapman and Harris 2004; Greany and Higham 2018). Choice and competition associated with these systems are against the equity and shared responsibility needed to enhance quality and equity in education systems (Sahlberg, 2015).
- In England, high-stakes consequences for those schools that are classified as failing are negatively affecting a proportion of the population that tends to be the most disadvantaged (Hutchinson 2016; Greany and Higham 2018).
- Studies comparing school inspection mechanisms within countries show differential effects partly depending on the overall judgement of the school’s effectiveness (Penninckx et al. 2014; Penninckx et al. 2016; Ehren and Shackleton 2016). Differential degrees of “accountability pressure” on school leaders (Altrichter and Kemethofer 2015, 32; Greany and Higham 2018) and teachers (Penninckx et al. 2016) have been reported.
- Low-performing/failing schools tend to receive greater levels of pressure, as well as embark in differing patterns of improvement after inspection (Matthews and Sammons 2004; Ehren and Shackleton 2016).

Judging schools as failing leads to...

## Improvement

- Speed recovery (Stark 2004; Mathews and Sammons 2004), acting as a catalyst for change (Allen and Burgess 2012) in 'just' failing schools
- Measure and Punish: need of tough penalties for those schools that do not comply (Amrein-Beardsly 2014)
- Stress inertia: inability to respond effectively needs to be overcome through external pressure and support (Jas and Skelcher 2005)
- Powerful symbolic function for policy and for the rest of the schools

## Spiral of decline

- Low teacher morale, feelings of impotence in these schools and the general public (Stoll and Myers 2002; Nicolaidou and Ainscow 2005)
- High levels of teacher mobility, stress and burnout (Perryman 2010)
- Weaker school ethos and morale (Jeffrey and Woods 1998; Elton and Male 2015)
- Narrow the curriculum to focus on tested subjects at the expense of other subjects and activities (Ehren and Shackleton 2016)



*Thanks for your attention!*

Dr. Bernardita Muñoz  
[b.munozchereau@ucl.ac.uk](mailto:b.munozchereau@ucl.ac.uk)