



What does it mean to widen participation? a machine learning analysis of historical and contemporary access discourses

Benjamin Hart¹

Received: 8 May 2025 / Accepted: 15 December 2025
© The Author(s) 2026

Abstract

Diversifying the demographic makeup of higher education institutions remains at the forefront of both government and university policy agendas. This paper examines the evolving discursive landscape of Widening Participation (WP) in the English higher education system through a machine learning analysis of institutional texts. Drawing on a novel dataset produced for this research entitled the Access Corpus, which is comprised of twenty years of Access Agreements and Access and Participation Plans, this paper applies machine learning tools to explore discursive shifts that have taken place in the domain of WP. Topic Modelling shows that both conceptual understandings of key terms and the *focus* of widening participation discourses have been transformed over the past twenty years as discourses have moved towards marketisation, and regulatory compliance. Importantly these shifts are not felt equally across the sector, statistical analysis indicates that high-status institutions display greater autonomy in their responses to regulatory policies and thus a more stable discourse over time. This paper's empirical findings are situated within a broader theoretical understanding of discourse which emphasises its inherently untethered nature, where meanings are constituted and reconstituted through use over time; aligning with Ball's 'policy as discourse' framework. Thus, 'widening participation' is neither fixed nor stable, but instead continually re-understood through institutional responses to shifting political, economic, and regulatory circumstances. This paper provides both methodological innovations through its use of topic modelling and a critical examination of how widening participation is constructed by higher education institutions over time.

Keywords Widening participation · Higher education · Fair access · Discourse · Policy · Enactment · Stratification

✉ Benjamin Hart
Benjamin.hart@education.ox.ac.uk

¹ Department of Education, University of Oxford, Oxford, Great Britain

Introduction

Widening Participation (WP) is a central policy concern in English higher education (HE) and reflects concerns about social inequalities in access to higher education. Despite concerted attention to narrow inequalities, access remains socially stratified with those from socioeconomically disadvantaged backgrounds and ethnic minority groups underrepresented within higher education. Whilst research has examined widening participation policies, much of this work has relied on either a close qualitative examination of texts, or on specific access issues (Boliver & Powell, 2023), often at particular moments in time (McCaig, 2015; Bowl & Hughes, 2013) or within a small number of institutions (Pickering, 2019). Thus, comparatively little attention has been placed on the discursive construction of widening participation by institutions over an extended period of time, or how these constructions vary by institutional type. The present paper addresses this lack of longitudinal insights into the construction of widening participation by examining how discourses have shifted over nearly two decades.

Drawing on Ball's (1993) conceptualisation of policy as discourse, Access Agreements (AAs) and Access and Participation Plans (APPs) are viewed as sites in which the meaning of widening participation is produced in response to wider and unstable political, economic, and regulatory conditions. The meaning of WP is problematised as inherently unstable, shifting in response to external pressures. Using topic modelling to analyse a large longitudinal corpus of institutional policy documents, this paper investigates how WP discourse has evolved over the past two decades, and whether institutions of different status respond differently to moments of policy change. The paper thus contributes to the existing literature both substantively and methodologically: first, by providing sector-wide longitudinal insights into an area which has not been examined in this way, and second, by demonstrating how computational approaches can be used to analyse policy discourse at a scale not easily achievable through many traditional approaches.

Background and literature review

Widening participation

The English HE system is weighted in favour of those from relatively privileged backgrounds (Chowdry et al., 2013) who are more likely to attend university (Younger et al., 2019). Indeed, a UCAS report (2022) found that school-leavers from neighbourhoods with the highest rates of HE participation are 2.4 times more likely to attend university than those from the lowest. This is widely recognised by HE institutions in England where the agenda to 'level the playing field' can be observed through the existence and organised activities of their Widening Participation (WP) departments. National funding and regulatory bodies have played a large role in WP, specifically Higher Education Funding Council for England (HEFCE) in the past and currently the Office for Students (OfS), where the latter requires individual HE institutions to set WP goals and to evaluate them effectively. The size of the field is illustrated by the fact that in 2017-18 £248 million was spent on WP (Robinson & Salvestrini, 2020).

WP is a broad term used to describe a number of processes and activities through which students from underrepresented groups are informed about the different benefits

of an HE experience and encouraged to attend HE, this can include activities designed to demystify university such as visit days, open days, student mentoring and summer schools. More recently, the evaluation of these WP activities is becoming increasingly emphasised (Hart, 2024).

Though early Access to Higher Education programmes were generally well received in education, in the 50+ years since their inception, the focus on *why* they are important seems to have shifted; this can, in large part, be explained by the increasing marketisation and neoliberalisation of English universities and the ascendancy of human capital theory (Becker, 1964). While earlier WP activity, since the 1970s and through to the early 2000s, often emphasised collaboration and social justice, the subsequent policy environment has increasingly foregrounded economic and competitive rationales. One of the key tenets of human capital theory is that education is a benefit for the individual who can expect a financial reward for their investment in pursuing education (Marginson, 2019). This framing has become progressively more visible in WP policy discourse over time and is often couched in terms of neoliberalisation of HE: “universities are now regarded in neoliberal policy as producers of knowledge and the products of knowledge that can enhance individual and collective human capital and so positively affect economic growth and development” (Patrick, 2013: 3).

A further, and more recent, development in WP has been the move away from earlier ‘aspiration raising’ discourses towards what Harrison (2018) terms a ‘possible selves’ framework. The emphasis on WP work moved more towards supporting students to imagine concrete and attainable futures for themselves and to see the pathways through which this might be achieved; in policy terms, this has meant a reduced reliance on aspirational rhetoric and a greater focus on activities that connect students’ present circumstances to actionable routes into their desired futures. More recent studies have suggested that the possible selves framework offers a useful way to engage learners by linking present actions to imagined futures (Jones et al., 2021).

HE institutions in the UK are not homogenous and attendance at one or another university has consequences for an individual’s progression to top professions and higher graduate salaries which vary significantly according to university attended (Belfield, 2018). Increasing participation to HE for those from under-represented groups needs to be considered alongside the issue of access to higher status institutions because of the implications of this for subsequent graduate level earnings and entry to higher status professions (Wakeling & Savage, 2015; Brooks et al., 2025). However, access to higher status universities for those from disadvantaged backgrounds is problematic (Boliver, 2013).

Higher education regulator

The regulation of the HE system has changed substantially over the past 20 years; this is evident in the change in regulatory body from the Office for Fair Access (OFFA) to the OfS. Unlike the more advisory OFFA, the OfS can sanction institutions based on its own criteria: the OfS can evaluate, monitor, and oversee institutional performance in several distinct domains and, based on its evaluation, monetarily punish institutions if they deem that the institutions are not performing to the requisite standard. Indeed, the ability to intervene in the HE landscape in such an overt way is clearly a marked shift in the *type* of regulatory atmosphere that prevailed when OFFA existed; there is a clear move from the *soft* power of OFFA, to the *hard* power of the OfS.

It has been argued that this ‘hard power’ that the OfS wields is potentially exploitable by the central government who have placed pressure (soft, or hard) on the regulatory body: rather than focussing on its core remit, the OfS has been criticised for being influenced by the shifting priorities of successive ministers, leading to the involvement in politically charged issues, for example freedom of speech, which fall outside of the regulators original purpose (Parliament, 2023). This is particularly alarming given that the OfS is intended to be an arm’s length body; a part of the machinery of government that is independent from ministers within the government, but still attached. Thus, the accusations that the OfS acts under the influence of ministers and Members of Parliament signals concern that the mission and aim of the OfS are under threat.

A recent House of Lords report further supports the concerns that the OfS is not acting as originally envisioned. The report is damning with regards to the impartiality of the OfS, as can be seen from the quotation: “The DfE’s [Department for Education] guidance to the OfS is often specific about what the Government wants the OfS to do, in contrast to other models this Committee has observed in its scrutiny of other regulators...by way of example, the Government’s guidance to the OfS on strategic priorities for the financial year 2022-23 explicitly says the OfS should include its approach to dealing with sexual harassment “in a condition of registration as soon as possible”” (House of Lords, 2023, p.78). This quotation, despite not concerning WP issues, effectively encapsulates the current state-of-play of HE regulation in the UK HE sector where the government is able to assert pressure on an institution which is designed as impartial and illustrates the problematic regulatory landscape which is clearly experiencing a moment of instability.

Two moments of flux in the higher education landscape

Two key changes have reshaped the HE sector over the past two decades, the first is the tripling of tuition fees in 2012 and the second is the introduction of the OfS in 2018. The first intensified market logics within the system with fees rising from £3000 to £9000 per year for undergraduate students. Despite this being an upper limit, universities adopted this maximum almost universally, driven by concerns that charging less would signal a lower quality offering. This episode illustrates the deepening influence of market logics in the sector, where tuition fees became perceived as proxies for institutional worth and value (Moore et al., 2011).

The tripling of fees may be seen as an extension of a broader trajectory of marketisation across the HE system (Brown & Carasso, 2013), shaped by earlier changes such as the introduction of tuition fees recommended in the Dearing Report (Dearing, 1997). In this context, degrees are increasingly viewed as mechanisms for individual economic advancement through increases in earning, employability, and access to high status jobs (Tomlinson, 2018). The framing of HE in this way is counter to conceptualisations of HE as a public good which emphasises personal development, critical thinking, and civic engagement (Marginson, 2011). The increasing marketisation of HE has reconfigured the relationship that students have with universities, which is viewed through a more transactional lens with students acting as consumers (Bunce et al., 2017).

The second moment of flux was the establishment of the OfS which appears to have further reinforced the trajectory of marketisation and regulatory control over the sector (Oliver, 2025). The consequences of this development have been profound: universities

are under more intense scrutiny to conform to a strict set of regulations concerning student access to, and participation in, HE. In relation to WP, institutions are necessitated to produce an APP when required by the OfS and this document must be approved by the regulator: if it is not approved, institutions may be subject to “punitive OfS sanctions” (McCaig, 2025, p. 24). These key shifts provide context for the present research and shape the discursive environment within which WP is constructed. Looking longitudinally at such moments of rupture is vital as it allows us to examine how the results of these moments of flux become embedded in the structure of discourse and shape the epistemic contours of evolving discourse of WP.

Policy as discourse

This paper uses ‘Policy as Discourse’ (Ball, 1993) as the basis of its theoretical framework. This means that policies are treated as discursive artefacts that *shape* and *are shaped* by regulatory structures and ideological currents. Importantly, they comprise an ever evolving ‘process’ by which meanings are structured and restructured in the production of discourse: “Discourses are about what can be said, and thought, but also about who can speak, when, where and with what authority” (Ball, 1993, p. 14). Policy is therefore not viewed as a fixed object with static and well-defined goals; meanings are produced and redefined through patterned linguistic practices.

As a result of this underlying theoretical lens, APPs are investigated as sites of discursive negotiation where widening participation is framed through the mediating forces of: the market; performance accountability; conceptualisations of social justice; amongst others. This allows for an examination of these documents as a space in which institutions outwardly perform to the regulator, (re)produce dominant logics, and construct the nature of ‘widening participation’. This theoretical underpinning determines the way in which one can investigate the documents that comprise the investigated corpus within this paper, since one should foreground both the linguistic patterns evident within it, as well as the broader policy landscape that allows for the emergence particular framings.

Through this lens, one of the key approaches to understanding policy as discourse is through an examination of these very linguistic patterns which shift over time, and which are evident in policy documents. This paper employs topic modelling for precisely this purpose: to investigate underlying linguistic patterns and, beyond simple word-counts, to identify latent semantic structures (Daenekindt & Huisman, 2020) which organise widening participation discourses. In analysing changes in topic prominence as an indicator of evolving discursive formations, this methodological approach operationalises Ball’s theory of policy as discourse; the quantitative patterns evident in the model provide empirical evidence of how the construction of ‘widening participation’ changes over time and in response to a wider political and regulatory landscape.

Understanding historic WP development through the access corpus

Given the complex, dynamic, and evolving nature of WP it is essential to examine how institutions themselves constitute and reconstitute the meaning of WP through their own documents. APPs and AAs provide a useful starting point from which to seek insights into institutional WP policies and plans.

Bowl and Hughes' (2013), in their analysis of eight AAs found that the most selective institutions were insulated from the direct pressures of government on them to widen participation; they “did not feel under pressure to take the ‘determined action’ called for...” (p. 22). The differentiation of universities on this front is a product of a quasi-marketised system in which access is most inequitable at elite institutions (Boliver, 2013). Furthermore, this mirrors the historically ‘merit-based’ approach taken by selective institutions in determining their undergraduate intake (McCaig & Adnett, 2009). This admissions approach has been under intense scrutiny more recently in the APP literature, Boliver and Powell (2023) suggest that the high status universities are increasingly progressive in their understanding of the issues which lead to access gaps: “there has been a shift in thinking away from interpreting lower attainment in terms of individual deficit and towards a more structural understanding of social inequalities” (p. 51). Further research involving AAs support these points, suggesting that post-1992 institutions are expected to make more changes in the face of policy, and that this same group of universities are expected to focus more on employment outcomes of students (McCaig, 2015).

While these studies provide useful insights, they often focus on small samples of institutions and years, and often on specific conceptual questions. The discursive construction of WP has remained somewhat neglected, specifically from a sector-wide and longitudinal approach; this is a particularly important area of research given the increase in regulatory pressure and competition within the sector, and the importance that key moments of flux have had on the field. It is important to see how the sector has developed over time so we can learn from the past to inform the future of WP. This paper contributes to the field by addressing a lack of longitudinal understandings of WP through the creation and analysis of a linguistic corpus, the Access Corpus, which is a novel dataset comprised of 20 years of institution texts and created for this research. By examining such a large amount of text directly pertaining to WP issues, and produced by institutions across the sector, this research offers novel insights into how WP has been conceptualised across time and across different institutional ‘types’; this research examines the impacts of changes to the HE landscape on institutions of different status levels. Informed by Ball’s (1993) conception of policy as discourse, the analysis treats these texts as active sites in which the meaning of ‘widening participation’ is produced and negotiated in response to shifting political, economic, and regulatory contexts.

The research questions of this paper are as follows:

1. How is ‘widening participation’ discursively constructed, at a sector level, by HE institutions in England?
2. How do institutional differences (i.e. levels of ‘status’) influence institutional level widening participation discourses?

Data & methods

The data: the access corpus

In order to answer the key research questions above, it is necessary to use data that reflect institutions, *cross-sector*, and to conduct analyses that can effectively capture meanings and

meaning shifts both over time and across institutions: we must capture institutional WP discourse. A good deal has been written about how to select a corpus of statements which relate to the delineation of an object of study. In addition, it should be noted that a shared object is vital for a corpus (in this case WP): “what binds a corpus of statements together in terms of its internal validity is that each statement forms the ‘conditions of possibility’ for the studied phenomenon” (Arribas-Ayllon & Walkerdine, 2008, p.12). The documents comprising the present corpus all concern the object *WP*. The documents examined as part of this research comprise institutional documents outlining individual WP and evaluation stances, specifically AAs and APPs.

HE institutions produce a, largely unstructured, document each year which signals their intentions and goals in relation to widening access to HE and carrying out outreach activity; these were initially called Access Agreements (2006–2019) and subsequently entitled Access and Participation Plans (2019–present). Although there are differences in the way in which these documents are structured, and the requirements for each have evolved over time as a result of changing policy priorities, the two documents are substantively comparable: the central policy purpose has remained over time, that of WP to HE for under-represented groups. These documents are publicly available on the OfS website and on the websites of each HE provider.

Institutional WP plans form the Access Corpus on which all of the analysis takes place. It consists of annual institutional WP documents from the past 20 years from 109 English HE institutions. Documents are included for the years 2005/2006 to 2024/2025, with the last document for each institution covering multiple years: 2020–21 to 2024–25. In sum, the Access Corpus contains ~12 million words ($n = 12,223,799$). Specialist institutions were discounted from the Access Corpus because of their specific characteristics which make them unique and thus less comparable.

Creation of the corpus

All AAs and APPs were downloaded from the OfS website for each of the 109 institutions accounted for in the Access Corpus. All documents were combined into a single file and pre-processed for analysis.

First, the text from each document was converted to lowercase in order to ensure lexeme uniformity (e.g., so “University” and “university” are treated the same). Second, any weblinks are removed from the documents since this would create noise in the data. The next key step in preprocessing is lemmatisation, where words are reduced to their lexemic root, for example “supporting”, “supported”, and “supports” all become the root, which is “support”; this means that morphologically similar variants of lexemes are not counted as distinct in further analysis. The corpus was then tokenised and transformed further into a document-term matrix, following the removal of standard stop-words such as “the”, “and”, “with” etc. The final step was removing rare words, thus words that were used by fewer than 5% of institutions were removed; this helps to focus the analysis and remove noise in the data (Feinerer et al., 2008). While this approach may exclude some highly specific policy references, it ensures that the resulting topics reflect patterns of language use that are more widely distributed across the corpus, improving stability and interpretability.

Methods

Latent Dirichlet Allocation (LDA) topic modelling is a machine learning tool which is used to determine latent word clusters in a linguistic corpus; this type of model is Bayesian and “each item of the collection is modelled as a finite mixture over an underlying set of topics. Each topic is, in turn, modelled as an infinite mixture over an underlying set of topic probabilities” (Blei et al., 2003, p.993) – these word clusters, or latent semantic structures, are referred to as ‘topics’. The chief benefit of this type of analysis is that it is not only unsupervised, but also can take into account researchers’ priors; this means that the analysis can compute vast quantities of data without constant input from the researcher, yet it is undertaken on the basis of *some* researcher input at the inception. Researcher input, in this context, refers to determining how many latent topics one would expect in the corpus – this is often determined by either (a) creating an elbow plot which shows the point of diminishing returns (i.e. where an extra topic in the model does not add in terms of explanatory value) or (b) coherence of topic words (this is a qualitative measure where the researcher assesses whether the topic words “make sense” based on a priori knowledge of the broad topic). Indeed, a combination of these two techniques informed the process by which seven topics were chosen (Grün & Hornik, 2011).

Topic modelling is used to identify discourse change over time by tracking the ways in which topics have changed throughout the corpus; this analysis can then be extended to explore sub-groups by explicitly exploring the patterns for particular groups of universities – in this case, levels of university ‘status’. A topic model can produce several useful outputs for discourse analysis, for example, identifying the most common words in each topic, and then listing words which are overrepresented in a particular topic; this information allows the researcher to understand more fully the context in which topics were discussed. In this scenario, ‘overrepresented’ refers to words that appear more frequently in a particular topic when compared to any of the other topics. A principal benefit of using topic modelling is that it does not rely on counting the frequency of isolated terms, instead identifying latent semantic structures by modelling how words co-occur across documents (Blei et al., 2003). This method allows for an analysis of changes in meanings within WP discourse, above and beyond simple lexical variation.

Boliver’s (2015) clustering of UK HE institutions was followed to determine the ‘status’ grouping of universities. In this scheme institutions were clustered based on teaching quality, research activity, economic resources, academic selectivity, and socioeconomic mix (Boliver, 2015). The results of this clustering analysis separate universities into four distinct ‘status’ groups, broadly speaking these can be described as (“n” refers to number of universities from this cluster in the Access Corpus):

- Cluster 1 – Oxbridge (Oxford and Cambridge) ($n=2$)
- Cluster 2 – Russell Group and other ‘old’ institutions ($n=38$).
- Cluster 3 – Most post-1992 institutions ($n=60$).
- Cluster 4 – Remaining post-1992 institutions ($n=15$).

The use of computational methods such as topic modelling and frequency analysis in combination seeks to track the shifting discursive constructions of widening participation over time. Informed by Ball’s (1993) theory of policy as discourse, the analysis views institutional

documents as sites of meaning construction and negotiation. Topic modelling allows for the identification of reoccurring patterns of language use which are interpreted to explore how particular framings emerge, and how regulatory changes shift language use – this constitutes a form of discourse analysis. In addition to topic modelling and associated output, this paper uses bigrams to complement the analysis; that is two words that cooccur across the corpus. This approach is increasingly used in HE policy studies either as a primary or secondary analytic technique (e.g. Cuesta-Delgado et al., 2025) which allows for an understanding of word frequencies over time. Bigram analysis supplements topic modelling given that the former investigates surface-level frequencies whereas the latter explores underlying latent semantic clusters of word uses (topics).

Through addressing these questions, the purpose of this paper is to contribute to the evidence base on the discursive construction of widening participation by analysing a large and presently underutilised dataset. The analytic approach allows for the identification of patterns and ruptures that are not visible in smaller or cross-sectional studies, and provides an empirical foundation for understanding the relationship between policy discourse and institutional stratification. This paper contributes methodologically by operationalising Ball's (1993) conceptualisation of policy as discourse through topic modelling, treating topics as latent semantic formations whose evolution over time reflects underlying discursive change. By applying this approach to institutional policy texts, the paper demonstrates how longitudinal corpora can be created and analysed systematically, allowing for the identification of discursive shifts over time that move beyond the scale achievable through many traditional approaches.

Whilst the primary analysis is exploratory in nature, it is partially informed by existing literature pertaining to marketisation and managerialism within HE (e.g. Brown & Carasso, 2013; Bunce et al., 2017). Based on this literature, it was anticipated that market-oriented and compliance-related language would increase over time, and particularly at key points of policy change: for example, at the tripling of tuition fees in 2012, and introduction of the OfS in 2018. Alongside providing evidence about these hypotheses, the analysis was also designed to detect unanticipated changes in discourse that might enhance our current understanding of how WP is constructed in policy texts.

Results

This section presents the results of the analysis conducted, which are structured in alignment with the research questions. First, the seven topics identified by the topic model are outlined and explored, which provides a backbone for subsequent analysis exploring how topic frequencies change over time. Following this, institutional differences are presented by breaking institutions into *types* (Boliver Clusters) and observing how different institutional discourses shift over time. It is important to note that the topics presented do not represent simple clusters of frequent words, but instead reflect underlying discursive constructions identified from analysing patterns of word co-occurrence. This provides insight into how the conceptual framing of WP shifts across time and institutions. The results presented in this section therefore represent a discursive formation that reflect how institutions frame and understand widening participation; subsequent sections provide an interpretation of each topic which links linguistic patterns to the wider political and regulatory landscape.

Topics

Table 1 represents (i) the top 10 words in each of the 7 topics throughout the entire corpus; (ii) the top 10 words in each of the topics that are “overrepresented” in the corpus. This “overrepresentation” is calculated by comparing the β value for a specific word in a specific topic with the average β for that word across all topics and dividing by the standard deviation of the β across all topics; thus, each word is given a positive or negative value based on its comparison to the mean (presented as a differential beta in standard deviations). Gap Analysis refers to analyses that universities conduct in relation to ‘gaps’; this could be defined as gaps in institutional performance in relation to student access, continuation, or progression. Of particular note are the words “Datum”, “POLAR” (post-code measure of participation in HE), “BAME” (Black and Minority Ethnic), “Black”, “White”, “Continuation” and “Reduce”. These words indicate that this topic is particularly concerned with investigating and analysing gaps in performance between students from different socio-demographic backgrounds. Both “Access”, “Attainment”, and “Continuation” are terms mentioned which signify the pervading nature of these “gaps” across the defined student lifecycle. The terms “Datum”, “Plan”, and “Rate” are likely to feature in the strategies that universities have devised to address these student performance gaps.

Funding represents discussions with respect to funding and the provision of money for students attending university; this can be inferred from use of the terms “Bursary”, “Tuition”, “Scholarship”, and “Fee”. The issue of student fees and funding arrangements have been much debated in WP circles and many HE leaders previously believed that providing direct monetary support to students wishing to enrol in HE would be an effective method of “doing access” (McCaig, 2010). Thus, it might reasonably be expected that this topic will appear more frequently at the beginning of the corpus, and gradually decline over time as views about the benefits of providing direct financial support to students changed – this pattern would be observed from subsequent analyses. This decline would illustrate how shifts in the funding regime, particularly after the 2012 tuition fee reforms, reoriented institutional discourse away from direct financial support towards other, more compliance-driven framings of access.

The topic ‘progression’ is defined by the OfS as both, in terms of moving on to *either* postgraduate study *or* employment, and also as generally understood *within* the HE sector as student success in moving through successive years of study and successfully completing a course. Variation in the prominence of this topic over time and between clusters may reflect differing institutional priorities, with some focusing on graduate outcomes as a marker of WP “success,” and others giving less discursive attention to progression, perhaps due to differing regulatory pressures or local student demographics.

Widening Access broadly represents “widening participation” strategies and, in particular, the specific events and forms of pupil engagement that universities facilitate for schools: “School”, “Activity”, “University”, “College”, “Outreach”, “Access”, “Programme”. There is also a specification concerning the *type* of event format, beneficiary or purpose: “Day”, “Pupil”, “Event”, “Session”, “Advice”, “Aspiration” (where aspiration is understood to involve fairly standard “aspiration raising” sessions for school pupils). Fluctuations in this topic’s frequency may signal changing attitudes towards the role of outreach events in meeting access goals, with potential declines pointing to a discursive shift from activity-based WP narratives towards outcomes-based or regulatory compliance framings.

Table 1 Most common words and overrepresented words by LDA topics

Topic 1	Topic 2		Topic 3		Topic 4		Topic 5		Topic 6		Topic 7		
Gap Analysis		Funding		Progression		Widening Access		School/College Links		Admissions		Access Performance Targets	
Most Common Words													
Term	β	Term	β	Term	β	Term	β	Term	β	Term	β	Term	β
Gap	0.026	University	0.030	University	0.021	School	0.025	School	0.018	University	0.025	University	0.026
Access	0.014	Fee	0.020	Access	0.014	Activity	0.017	Programme	0.013	School	0.020	Target	0.021
Datum	0.014	Bursary	0.018	Target	0.010	University	0.015	Access	0.013	College	0.014	Access	0.016
Attainment	0.013	College	0.013	Success	0.009	College	0.014	College	0.012	Target	0.012	School	0.014
Plan	0.012	Access	0.012	Progression	0.009	Outreach	0.012	Activity	0.012	Activity	0.012	Activity	0.013
Participation	0.011	Participation	0.011	Activity	0.008	Target	0.011	Target	0.009	Access	0.012	Outreach	0.013
Target	0.010	Income	0.010	Learn	0.008	Access	0.011	Education	0.008	Programme	0.011	Agreement	0.010
Continuation	0.009	Time	0.010	Provide	0.008	Participation	0.009	Participation	0.008	Outreach	0.011	Retention	0.010
BAME	0.009	Agreement	0.010	Study	0.008	Programme	0.008	Outreach	0.007	Provide	0.007	Programme	0.009
Rate	0.009	Education	0.009	Time	0.008	Fee	0.007	Study	0.007	Admission	0.007	Entrant	0.009
Overrepresented Words													
Term	$\Delta\beta$	Term	$\Delta\beta$	Term	$\Delta\beta$	Term	$\Delta\beta$	Term	$\Delta\beta$	Term	$\Delta\beta$	Term	$\Delta\beta$
BAME	5.734	Scholarship	3.292	Engagement	1.903	Day	1.663	London	2.261	Admission	2.396	Retention	2.688
Gap	5.540	Tuition	2.408	Graduate	1.811	Pupil	1.058	Pre	1.220	Teacher	1.50	HESA	2.372
White	4.609	Bursary	2.163	Success	1.609	Event	0.987	Session	1.038	Participant	1.449	Column	2.191
Continuation	3.780	Fee	1.956	Learn	1.521	WP	0.953	Project	0.809	Summer	1.233	Benchmark	1.986
Reduce	3.636	Income	1.665	Team	1.504	Session	0.941	Teacher	0.704	UK	1.070	Entrant	1.767
Black	3.587	Cost	1.621	Employment	1.454	School	0.928	Background	0.644	Cohort	1.005	Table	1.635
Compare	3.292	Additional	1.457	Practice	1.425	Advice	0.909	OFFA	0.596	Entry	0.996	Description	1.546
POLAR	3.287	Fund	1.119	Enhance	1.421	Aspiration	0.906	Scheme	0.593	Fair	0.991	Indicator	1.477
Attainment	2.953	Receive	1.093	Equality	1.363	London	0.884	Lead	0.577	Potential	0.866	Milestone	1.369
Investment	2.474	Scheme	1.066	Approach	1.322	Potential	0.770	Pupil	0.553	Subject	0.863	Collaborative	1.271

Note: β = Probability of a term occurring in a particular topic (word β s sum up to 1 within each topic); $\Delta\beta$ (Differential Beta) = $(\beta - \bar{\beta})/\bar{\beta}$

These words in the topic Schools/College (this refers to further education colleges which are often attended by 16–18 year olds to complete qualifications such as A Levels). Links refer to the links that universities have with school and colleges. Terms suggesting this include: “School”, “College”, “Target”, “Outreach”, “Session”, “Teacher”; there is even a geographic location in the “over-represented” word list, namely “London”, which further suggests that this topic relates to the kinds of links that universities have with school and colleges because of its dense pre-university population. A declining trend here could indicate a reduced emphasis on cultivating sustained partnerships with feeder institutions, potentially as resources and discursive space are redirected towards meeting regulator-defined performance targets.

The topic Admissions is signified by the words: “University”, “Admission”, “Cohort”, “Entry”, “Fair”, “Potential”. These words mirror those that are employed by institutions in relation to their admissions policies, for example, references to “fair access” or considering students’ “Potential” given appropriate contextual information about their backgrounds. Admissions discourse is likely to be most salient among higher-status institutions facing ongoing scrutiny over their selection processes, reflecting both external pressures to demonstrate fairness and internal imperatives to maintain prestige.

The most common and overrepresented words in Access Performance Targets are: “Target”, “Access”, “Retention”, “Benchmark”, “Indicator”, “Milestone”, and even “Table” which reflect how targets are reported and discussed. The growth of this topic over time would underscore the rising influence of performance management regimes in WP policy, marking a shift towards quantifiable accountability and away from more normative or values-based articulations of widening participation.

Change over time

Figure 1 illustrates how the frequency with which topics are spoken about changes over time, with each topic being a line on the graph.

Gap Analysis appears to be fairly stable over the corpus until 2018 when it rises dramatically up until the present day. Funding appears to be a popular topic within university output from the start of the corpus up to 2012 when there is a substantial decrease and then a plateau which remains the case to the end of the corpus. There is some fluctuation in the extent to which universities discuss the topic Progression, however, the overall trend is a positive one which increases in frequency up until 2012 when it declines significantly, followed by a steady rise to 2019, and then a sharp decrease in 2020. The overall trend for Widening Access is that of decline, with 2020 marking the year when this topic is discussed least. Indeed, there is a sharp decrease in institutional discussion of this topic from 2012 to 2013. School/College Links follows a similar pattern to Widening Access, following a generally downward trajectory from 2006 to 2020. The trajectory of the Admissions topic in the corpus follows the same broad downward trajectory, with a similarly large drop around 2012–2013 evident, followed by a steady decline from 2013 to 2020. The Access Performance Targets topic has a different trajectory compared to each of the other topics since it is relatively stable from 2006 to 2012, and then rises rapidly (in opposition to most other topics) to its highest point in 2013, followed by a slow decline to 2020 where the topic is less frequently discussed compared to 2006.

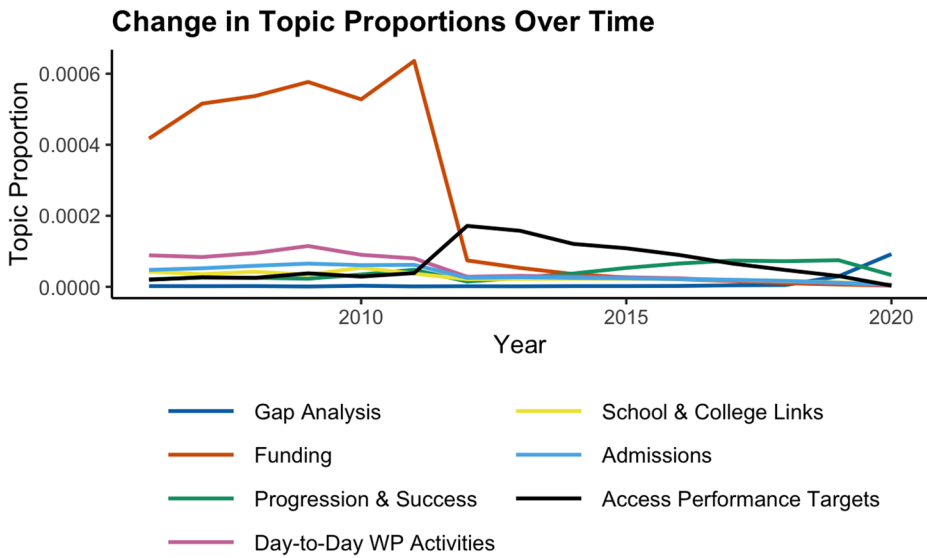


Fig. 1 Change in Topic Proportions Over Time

Figures 2, 3, 4, 5, 6, 7 and 8 represent disaggregated line graphs which represent change in topic proportions over time, broken down by Cluster with Cluster 1 representing the highest status institutions and Cluster 4 representing the lowest.

Figure 2 shows the proportion of the Gap Analysis topic over time. This remains relatively stable from 2006 until 2018, when there is a sharp and sustained increase across most institutional clusters. The timing aligns with the introduction of the OfS and its emphasis on evaluating attainment, continuation, and progression gaps, suggesting that the regulator's priorities directly shaped the language used in institutional plans.

Figure 3 displays the proportion of the Funding topic over time. References to funding are prominent in the earlier years of the corpus but decline sharply after the 2012 tuition fee reforms, before stabilising at a much lower level. This drop is consistent with a shift away from direct financial framing of access towards other approaches, reflecting both changes in sector funding structures and evolving assumptions about what constitutes effective WP practice.

Figure 4 shows the proportion of the Progression topic over time. The trend rises steadily from 2006 to 2012, dips slightly in the immediate aftermath of the tuition fee reforms, then recovers and peaks around 2018 before falling again in the most recent years. This pattern suggests that progression (both in the sense of course completion and post-graduation outcomes) was an enduring priority, but its prominence was periodically reshaped by major policy and regulatory changes.

Figure 5 displays the proportion of the Widening Access topic over time. There is a steady decline across the corpus, with the most pronounced drop occurring between 2012 and 2013. This suggests a move away from activity-based outreach narratives towards more outcome-focussed and compliance-oriented framings, particularly following the tuition fee reforms. Cluster 1 appears consistently low throughout the entirety of the corpus.

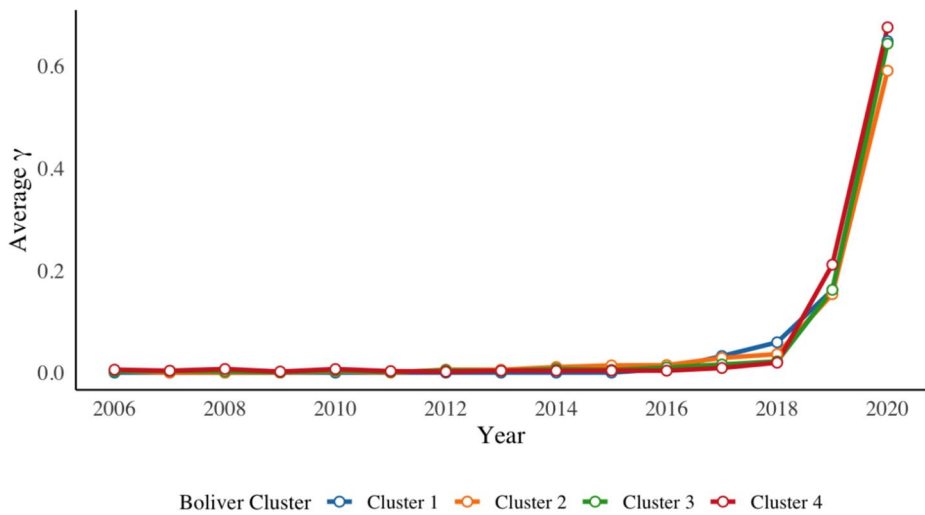


Fig. 2 Change in Gap Analysis Proportions Over Time

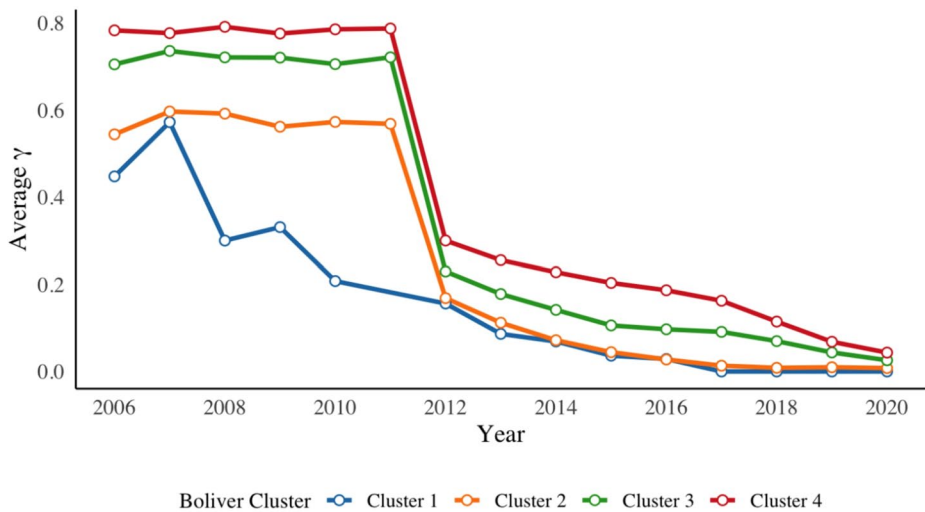


Fig. 3 Change in Funding Proportions Over Time

Figure 6 shows the proportion of the School / College Links topic over time. The trend is broadly downward. This indicates that explicit references to institutional partnerships with schools and colleges became less prominent, possibly reflecting a shift in discursive and operational priorities towards meeting centrally defined targets. Cluster 2 appears distinct from the other clusters in its pattern: it is consistently higher throughout the corpus and decreases significantly from 2018 to the end of the corpus.

Figure 7 presents the proportion of the Admissions topic over time. The proportion is relatively consistent in the earlier years (and high, particularly for Cluster 1) but declines,

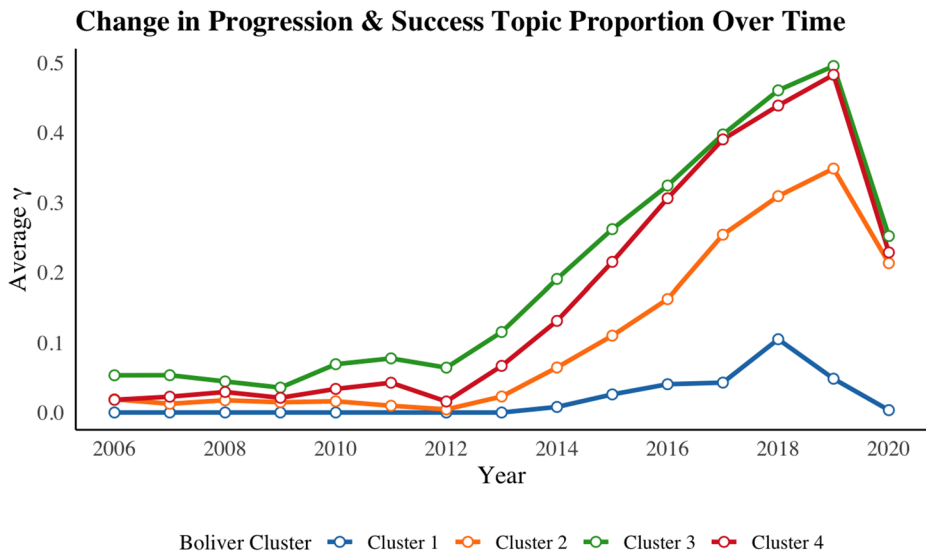


Fig. 4 Change in Progression Proportions Over Time

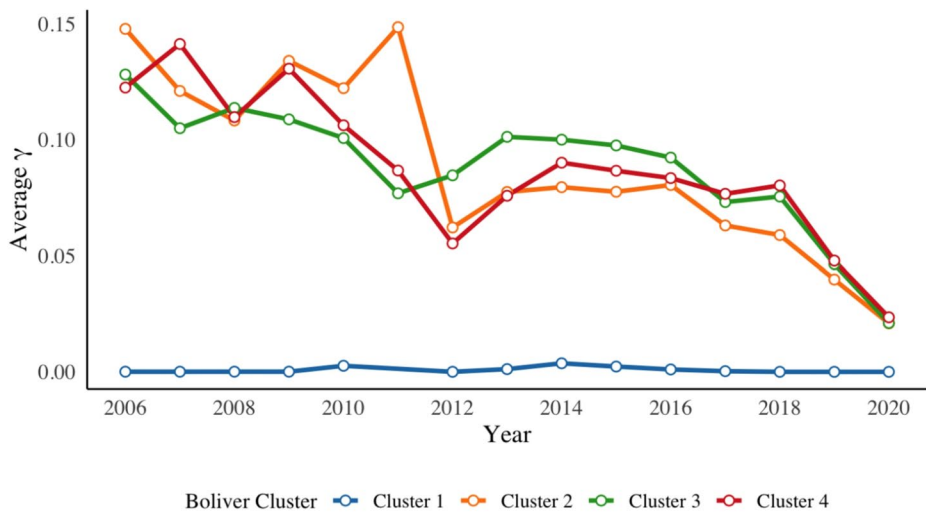


Fig. 5 Change in Widening Access Proportions Over Time

stabilising at a lower level post-2018. This suggests that discussions of admissions processes and fairness became less central to WP discourse, with institutions perhaps placing greater emphasis on retention, performance, and compliance narratives. Notably, Oxbridge (cluster 1) remains persistently the highest in their attention to admissions throughout the period. This sustained emphasis may reflect the intense public and regulatory scrutiny surrounding their admissions processes, alongside the highly competitive entry requirements that characterise these universities.

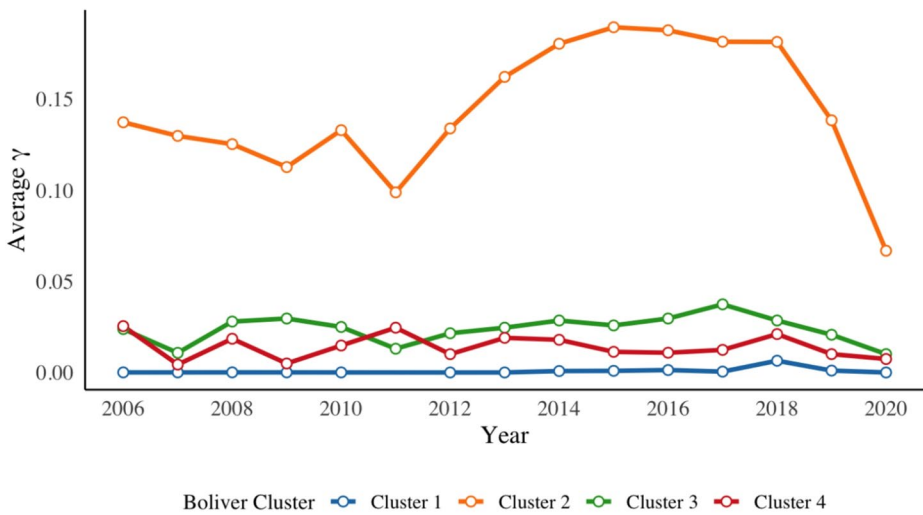


Fig. 6 Change in School/College Links Proportions Over Time

Change in Admissions Proportions Over Time

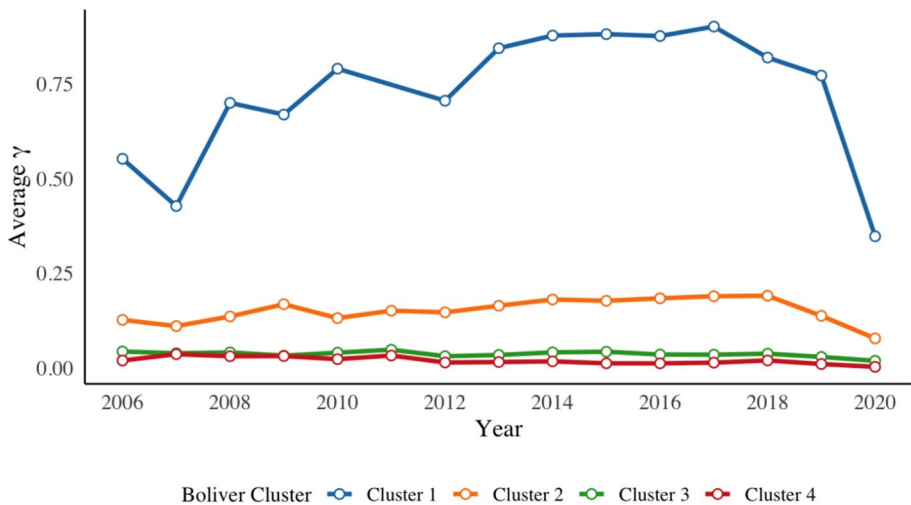


Fig. 7 Change in Admissions Proportions Over Time

Figure 8 shows the proportion of the Access Performance Targets topic over time. There is a sharp rise between 2011 and 2013, followed by a gradual decline, though remaining above pre-2012 levels – though this is not the case for Cluster 1, which does rise around 2012 but not to the same magnitude as the other Clusters. This spike corresponds with increased regulatory demands for measurable benchmarks, signalling the consolidation of performance monitoring as a dominant theme in institutional WP discourse.

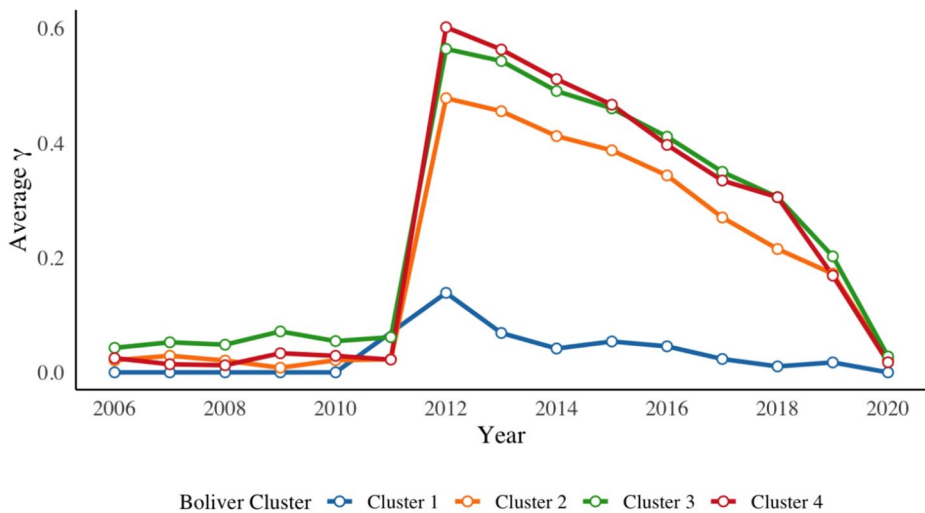


Fig. 8 Changes in Access Performance Target Proportions Over Time

Figure 9 tracks the frequency of the bigram “Competitor Average” over time (for combined clusters). It is almost absent in the early years but increases sharply to 2012, from where it remains in an elevated position, reflecting the embedding of market comparison language within institutional plans. This aligns with the broader marketisation of the HE sector post-2012.

Discussion

The findings of this paper explicate the seven topics that were spoken about by English universities over the years (2006–2025) of the Access Corpus, and disaggregated these findings by university status. Two moments of rupture were evident across the corpus which substantially changed WP discourse across the sector, with particular institutional status levels more impacted than others. Whilst the analysis highlights apparent temporal alignments between discursive shifts and major policy changes (notably the 2012 tuition fee reforms and the establishment of the OfS in 2018), these relationships should not be understood as singular or deterministic. Changes in topic prominence may also reflect shifts in institutional guidance and the templates provided for AAs and APPs, which could steer universities towards or away from particular framings (though this is still discursively important). The broad decline in the proportions of most topics after 2012 is likely a function of both regulatory change and structural alterations to the documents themselves. These patterns suggest that the observed declines are better understood as discursive re-weightings rather than an absolute reduction in policy attention. Although this study cannot isolate the precise contribution of template changes, the persistence and scale of the observed shifts (across all topics) indicate that they cannot be reduced to structural artefacts alone. If template changes were the primary driver of patterns evident in the data, we would expect topic proportions to stabilise once institutions adapted to the new format, instead what we see is that the altered discursive

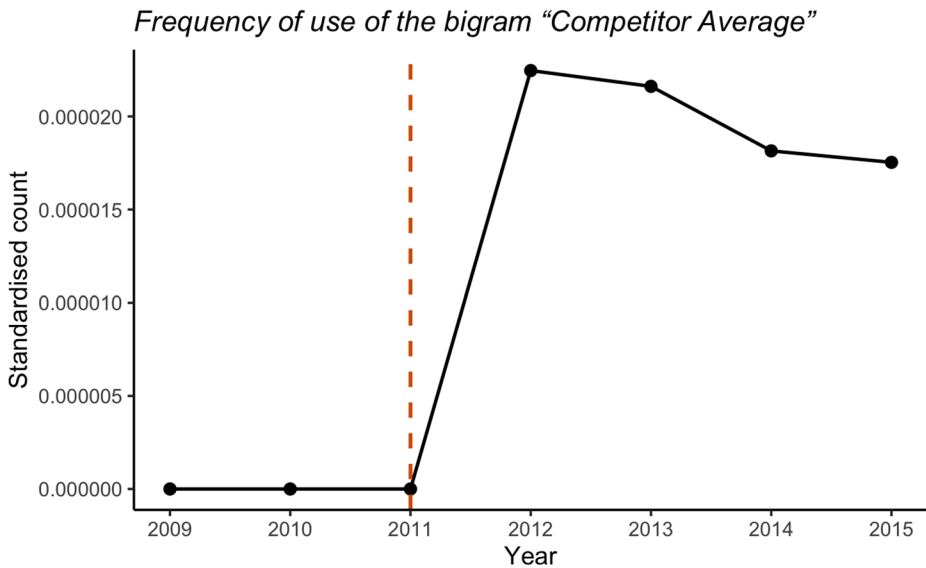


Fig. 9 Market of use of the bigram "Competitor Average"

sive patterns persist – this is evident across all institutional types. Furthermore, although framing and guidance for APPs evolved over time, there was no singular mandated template imposed that would produce a uniform structural shift in the corpus, thus any changes in emphasis uncovered in the topic modelling carry discursive significance.

Tripling of tuition fees

The tripling of tuition fees in 2012 to £9000 can be viewed as the cause of a rupture in WP discourses. This fee increase had a profound effect on the HE system, shaping both the actions of institutions and students; its consequences were wide-spread, impacting on institutional governance, the HE 'market', and the framing of WP.

The topic modelling analysis indicates that there are notable shifts in six of the seven identified topics around the time of the 2012 reforms in: Funding, Progression, Widening Access, Schools and College Links, Admissions, and Access Performance Targets. These shifts were not momentary deviations and they represent new discursive trajectories that persisted well beyond the 2012 fee increase. Indeed, topic frequencies do not return to their pre-2012 levels which suggest that a more transformational shift occurred in institutional governance and in the articulation and conceptualisation of WP work. Table 1 captures these discursive shifts; with the most striking trajectory being the topic Funding. Prior to 2012 Funding had been a dominant focus, but its frequency significantly decreases post 2012 and never recovers. The only topic to increase in the wake of the funding reform is Access Performance Targets.

It is important to emphasise that, for example, the decline in the Funding topic does not imply that funding ceased to be materially important to widening participation. It can instead be interpreted as a discursive reconfiguration where, post-2012, institutions decreasingly focussed on financial support and increasingly focussed on other policy priorities

(such as theories of change, evidence of impact, meeting targets). Thus, this particular shift signals a relative de-emphasis on funding within APP discourse as compared to other topics: the findings concern the organisation of discourse rather than the underlying material significance of funding.

One of the key transformations more broadly in the HE sector during the time period covered by the corpus is an increasing sense of marketisation permeating the system; this is succinctly explored by drawing from the work of Fromm (1978): “the current higher education (HE) market discourse promotes a mode of existence, where students seek to ‘have a degree’ rather than ‘be learners’” (Molesworth et al., 2009, p.277). The view of universities as businesses and students as consumers became more prevalent (Marginson, 2011; Bunce et al., 2017) and shaped not only the nature of education offered, but also language used by institutions. WP was not immune to this shift and it became increasingly common for universities access teams to be situated within student recruitment departments. While this marketisation is clearly articulated in the literature (Bendixen & Jacobsen, 2017; Teixeira & Dill, 2011; Brown & Carasso, 2013), it is also evident in the Access Corpus through the increasing embeddedness of language which reflects marketised governance structures. Figures 9 and 10 show the increase in bigrams which represent competition within the sector, and the intensification of this after the fee rise.

The rise of the topic Access Performance Targets topic post-2012 further supports this interpretation. This topic contains the words “target”, “benchmark”, “indicator”, and “milestone”, which are lexemes strongly associated with performance management. They suggest a shift in the institutional discourse of WP from one which was primarily normative and a ‘good’, to one which is increasingly viewed as something to be monitored, measured, and reported. As Brown (2015) notes, this language is symptomatic of the remodelling of university governance on corporate lines.

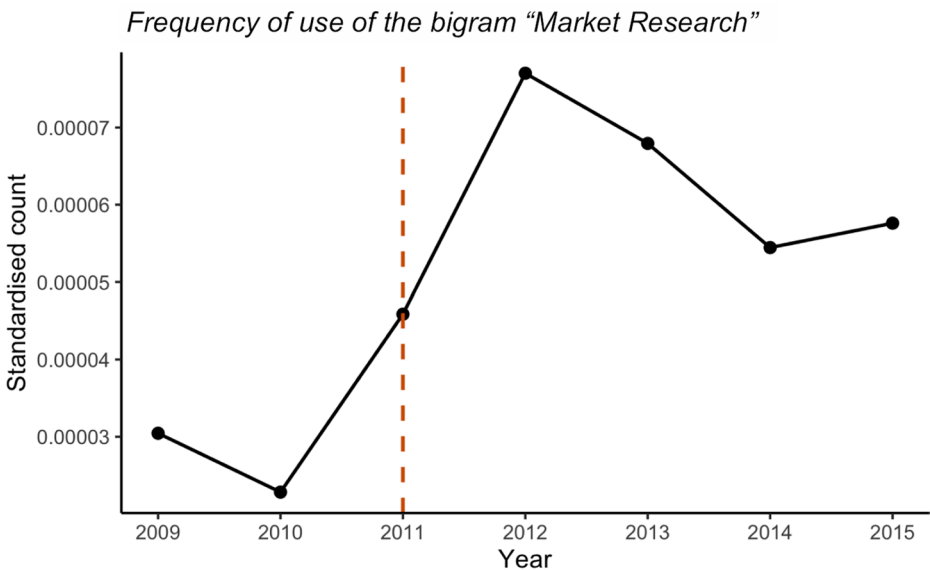


Fig. 10 Market of use of the bigram "Market Research"

The results thus show that the 2012 tuition fee tripling not only shaped the financial structure of HE in England but also significantly reshaped discourses pertaining to WP and its conceptualisation. After 2012 institutional discourse moved towards more technocratic and market vocabulary which reflects broader transformations across the sector.

Introduction of the office for students

As outlined earlier the establishment of the OfS in 2018 marked the introduction of a new regulatory structure for HE in England. The OfS has considerably more power over institutions than its predecessors, and the results of the topic model suggest that their introduction marks another rupture in WP discourse across the sector.

One of the most notable shifts post-2018 is the sharp rise in the prominence of Gap Analysis. This topic refers to institutional evaluations of differences in students' outcomes (access, continuation, attainment, progression) by background. Given the increase in frequency of this topic at this point in time, it can be inferred that this reflects the regulator's emphasis on evidence, evaluation, and measuring impact.

There is also a significant increase in the use of regulatory and compliance related language across the Access Corpus, post-2018. As can be seen in Figs. 11, 12, 13, and 14 words such as "report", "accountability", "declare", and "require" significantly increase over this period. This suggests that institutions began to frame their WP activities in language that mirrors OfS expectations, perhaps due to the regulator's powers of reprimand (House of Lords, 2023).

Ball (2003) critiques the increase of a culture of performance monitoring in education settings by noting the dangerous shift from first order, to second order activities. Where, in WP, first order might refer to going into schools and meeting students, and second order

Frequency of use of the word "report"

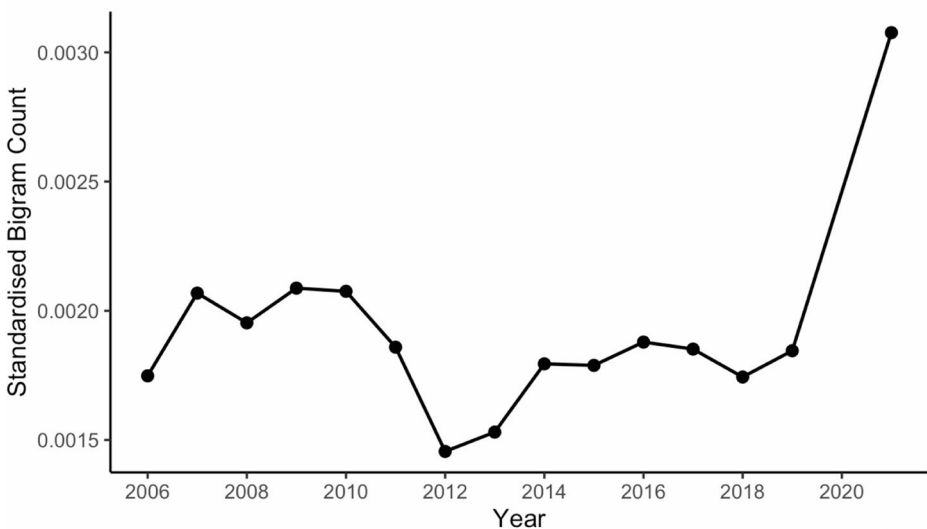


Fig. 11 Frequency of the use "report"

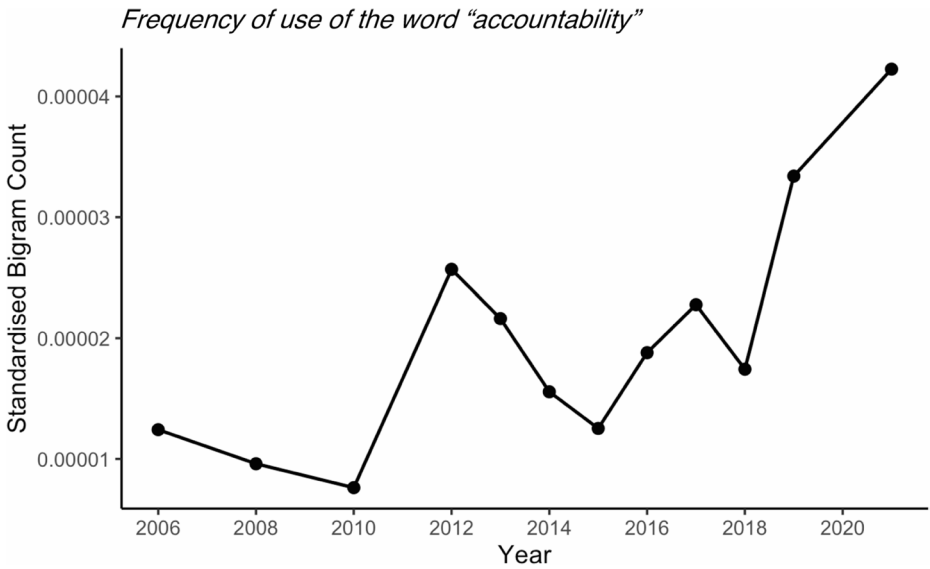


Fig. 12 Frequency of the use "accountability"

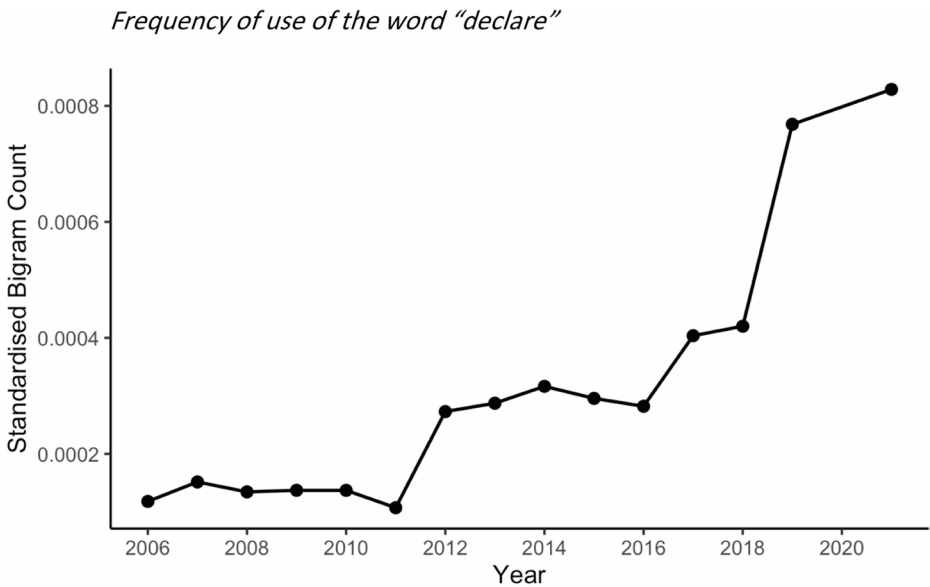


Fig. 13 Frequency of the use "declare"

might refer to institutional reporting and self-justification. It was noted earlier the dangers of over regulation in HE and mission creep of regulatory bodies, Lyotard notes: “the idea (or ideology) of perfect control over a system, which is supported to improve its performance, is inconsistent with respect to the law of contradiction: it in fact lowers the performance level it claims to raise” (Lyotard, 1984, p.55).

Frequency of use of the word "require"

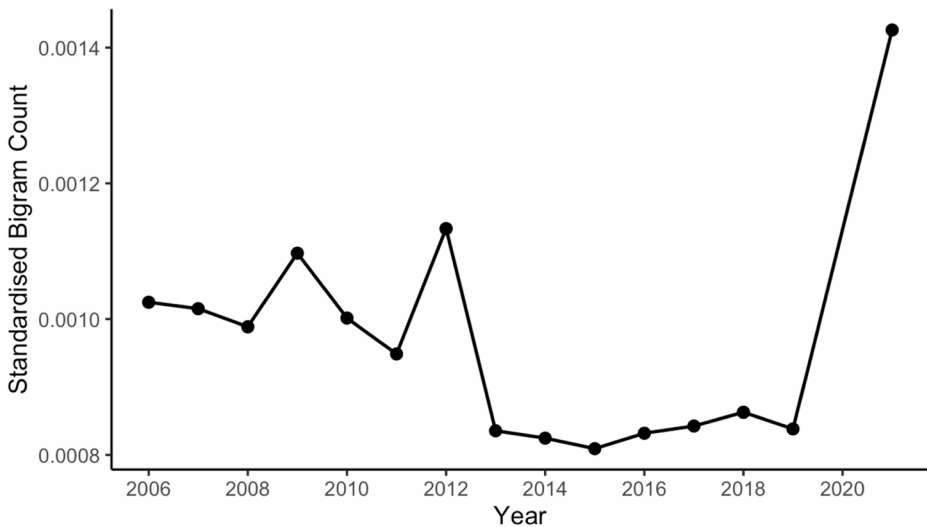


Fig. 14 Frequency of the use "require"

Together, the rise in Gap Analysis and the increase in regulatory language across the Access Corpus after the introduction of the OfS signifies another rupture in the discourse of WP, this time with the consequence of a foregrounding of evaluation, compliance, and quantifiable performance.

Institutional differences

This section addresses the second research question pertaining to institutional differences in WP discourses. First, relating to how different types of university emphasise different discourses relating to WP, and second, examining institutional reactivity to ruptures – specifically the two ruptures foregrounded in this paper, in 2012 and 2018.

The most prominent topic for Cluster 1 (Oxford and Cambridge) was Admissions; this is by far the most widely referenced topic, substantially higher than for any other topic from the LDA topic model. Given that the two universities that comprise Cluster 1 are the universities with the highest entry tariff requirements, it is unsurprising that the topic with the most references is Admissions. Entry grade requirements are the most challenging at these institutions and, it can be argued, this creates a very high admissions threshold which works against the circumstances of students who are from the most disadvantaged backgrounds (Boliver, 2013). In this situation, widening access to the institution for these students will necessarily be constrained by the stringency of the admissions policy (Kandiko Howson et al., 2022).

It is particularly notable that the standard admissions routes at these two institutions for almost all courses includes interviews, and additional subject-specific tests. Many schools across the UK have Oxbridge specific preparation for their students (Donnelly, 2014) – this is unique to Cluster 1 and is not widely found for other universities. Therefore, the admis-

sions criteria and attendant processes are markedly (almost uniquely) different for Oxford and Cambridge. In the context of widening access to HE for students from disadvantaged backgrounds, this operational feature may explain the significantly higher frequency of discussions relating to admissions by these institutions.

The unique circumstances and popular mythology surrounding Oxbridge institutions further manifests in other ways and beyond the key focus of their institutional APPs being placed on Admissions issues; we can observe, for example, that in the case of the topic Access Performance Targets (Topic 7), these institutions do not follow the same trend as Clusters 2, 3, and 4. More specifically, all institutions appear to be directly affected by the policy drive of 2012 which involved the tripling of tuition fees. This development brought with it an increased focus on performance targets and a closer monitoring of access activities, and yet the impact of this does not appear to have been as keenly felt by the Universities of Oxford and Cambridge. Though there is clearly an increase even for Cluster 1 in the wake of the 2012 policy drives, Cluster 1 is affected to nowhere near the same magnitude as any other.

Cluster 2 institutions also focus on Admissions but they display a comparatively greater emphasis on school and college links; this represents a dual commitment to maintain prestige whilst also demonstrating commitment to WP agendas. McCaig (2015) argues that Russell Group universities were incentivised by market structures to attract the brightest and best students for their own institutions, and this interpretation can be carried through to findings in the present paper which show that Cluster 2 institutions place a strong focus on school and colleges links. The topic Admissions is referenced the second most frequently for Cluster 2 (compared to other clusters); this may at least in part be explained by the fact that these institutions are some of the most prestigious in England and are thus under particular media, public, and policy scrutiny to make sure they are acting in the ‘correct’ way. This situation is similar to that found in Cluster 1.

Cluster 3 and 4 institutions demonstrate a shift in focus with particular dominant topics being: Access Performance Targets, Progression, and Gap Analysis. These topics represent an approach to WP which is mediated by regulation and technocratic agendas, highlighting the need for measurable outcomes. These discourses reflect the policy agenda of the HE regulator which focusses on accountability and evidencing impact.

The findings also align with earlier observations by McCaig (2015), who, through a discourse analysis of AAs, identified a shift in institutional policy discourses towards less favourable conditions for WP (particularly among post-1992 institutions) and described the ‘raising aspirations’ agenda as becoming a more difficult project. The present analysis extends this work by adopting a longitudinal, corpus-based approach covering the entire English HE sector, enabling the identification of how such shifts unfolded over nearly two decades and at key policy rupture points. This broader empirical base confirms McCaig’s proposition while also elaborating on the differentiated trajectories by institutional status, showing in detail how compliance-oriented discourses became more embedded in lower-status institutions over time.

Reactivity to policy by institutional status

The analysis shows that institutional discourses shift in response to ruptures: namely the 2012 tuition fee increase, and the 2018 establishment of the OfS. What is particularly interesting is that these ruptures are not felt equally across all institutions.

Following the 2012 tuition fee increase, most institutions changed their language in relation to WP and this is illustrated in a marked change in many of the topic model topics. However, the magnitude of these changes appears to drastically vary by Cluster. Institutions in Clusters 3 and 4 showed the most pronounced changes in their language use, particularly in relation to the topics Access Performance Targets and Progression, which indicate a reactivity to the policy agendas around that time. Cluster 1, on the other hand, appears to show only marginal differences in their institutional discourse around this point of rupture. A similar pattern is found in the 2018 rupture, with the most notable discursive shifts occurring in the topic Gap Analysis. This topic encompasses the use of evaluative frameworks to measure differences in access and progression for students from different backgrounds and was particularly pushed by the OfS once introduced.

When comparing the discursive trajectories of different Clusters across the Access Corpus, it becomes clear that there is a linear pattern evident which reflects institutional status stratification: with Cluster 1 being followed by 2 which is, in turn followed by 3, and then 4. A notable difference in this linearity appears when observing the trajectories of Cluster 1 institutions; these universities seem to chart their own course and thus appear more autonomous in the face of regulatory, economic, and financial pressures. This is supported by literature which notes that elite universities possess capitals which allow them to be protected from certain aspects of commodification, whereas “subordinate universities...are more likely to be buffeted by market forces” (Naidoo, 2004, p.471). While the distinctiveness of Cluster 1 (Oxbridge) undoubtedly contributes to the magnitude of the differences observed, similar directional trends are evident between Clusters 2–4, suggesting that the broader stratification pattern identified here is not solely driven by Oxbridge’s position. Rather, Oxbridge represents an extreme case within a wider status gradient across the sector.

This paper extends these earlier insights by providing a longitudinal, corpus-based account of how stratification is manifest in the language of policy texts over an extended period. Whilst prior work (e.g. Naidoo, 2004) has established that lower-status institutions are more responsive to regulatory demands, the present analysis shows how this responsiveness is realised through specific changes in language and topic emphasis following major policy ruptures. This analysis confirms the existence of differentiated institutional strategies and outlines the linguistic mechanisms through which compliance is signalled, offering a new empirical insight into the relationship between institutional status and regulatory practices.

Conclusion

This paper explored how WP is discursively constructed within institutional texts by drawing on machine learning tools to examine longitudinal and stratified patterns within a novel corpus, the Access Corpus. The findings show that institutional discourses pertaining to WP are dynamically constructed through complex negotiation within circumstances that are dictated by shifting regulatory, political, and economic conditions, but also particularly by *institutional status*. Seven topics were identified and explored in terms of their longitudinal trends, and analysis showed that the tripling of tuition fees in 2012, and the introduction of the OfS in 2018 acted as moments of discursive rupture in the HE system; these are important moments where concepts are redefined and there are opportunities for alternative futures to emerge; discourse is not only a reflective force, but also a constitutive one.

In line with Ball's (1993) theory of policy as discourse, the analysis presented in this paper reveals a plurality of institutional responses to external conditions. Though policy documents may appear to mandate a uniform response from disparate institutions, their interpretation is shaped by institutional characteristics, including material and symbolic capital. It is clear from the analysis presented in this paper that high status institutions are able to exhibit greater autonomy over their institutional discourses, whereas lower status institutions more easily succumb to policy mandates; this highlights unequal policy consequences across the sector, where particular institutions are able to: exert their own will, engage selectively, and ultimately resist particular policy agendas.

While the finding that WP is not a neutral or technocratic process is well established in the literature, this paper contributes to the literature by showing how such dynamics 'play out' over an extended time period. By combining longitudinal analysis with institutional stratification, the analysis conducted in this paper demonstrates that ruptures in the policy environment do not generate *uniform* change across all institutions, but instead change that is mediated by institutional status. Providing evidence to support this point is vital and highlights the limits of regulatory models that assume a level playing field in institutional capacity to adapt to policy changes. The argument here is not that some institutions ought to be more tightly regulated than others, but that policy frameworks need to recognise that institutions occupy unequal structural positions and therefore engage with regulatory priorities in uneven ways. Thus, this paper moves beyond re-stating the point that WP is embedded in unequal structures, but detailing how these inequalities are articulated in policy discourses, and how they shift in response to regulatory pressures. Future research might build on this work by examining whether these discursive shifts correspond to substantive changes in institutional practice and the delivery of WP activities, including how the move from 'aspiration raising' to 'possible selves' frameworks is reflected in policy discourse and if their underlying logics align with the increasingly marketised higher education space.

Funding No funds, grants, or other support was received.

The authors have no relevant financial or non-financial interests to disclose.

Ethics approval was given by the University of Oxford's ethics committee: Central University Research Ethics Committee Approval Reference: CIA-21-282.

Financial interest The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Arribas-Ayllon, M., & Walkerdine, V. (2008). Foucauldian discourse analysis. In C. Willig, & W. Stainton-Rogers (Eds.), *The Sage handbook of qualitative research in psychology* (pp. 91–108). Sage.

- Ball, S. (1993). What is policy? Texts, trajectories and toolboxes. *Discourse (Abingdon, England)*, 13(2), 10–17. <https://doi.org/10.1080/0159630930130203>
- Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18(2), 215–228. <https://doi.org/10.1080/0268093022000043065>
- Becker, G. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. National Bureau of Economic Research.
- Belfield, C. (2018). *Relative labour market returns to different degrees* (Research Report RR787).
- Bendixen, C., & Jacobsen, J. C. (2017). Nullifying quality. *Quality in Higher Education*, 23(1), 20–34. <https://doi.org/10.1080/13538322.2017.1294406>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993–1022.
- Boliver, V. (2013). How fair is access to more prestigious UK universities? *British Journal of Sociology*, 64(2), 344–364. <https://doi.org/10.1111/1468-4446.12021>
- Boliver, V. (2015). Are there distinctive clusters of higher and lower status universities in the UK? *Oxford Review of Education*, 41(5), 608–627. <https://doi.org/10.1080/03054985.2015.1082905>
- Boliver, V., & Powell, M. (2023). Rethinking merit? The development of more progressive approaches to university admissions in England. *Widening Participation and Lifelong Learning*, 24(3), 33–55. <https://doi.org/10.5456/wpll.24.3.33>
- Bowl, M., & Hughes, J. (2013). Discourses of fair access in English higher education: What do institutional statements tell us about university stratification and market positioning? *Widening Participation and Lifelong Learning*, 15(4), 7–25. <https://doi.org/10.5456/WPLL.15.4.7>
- Brooks, R., Hart, B., Henseke, G., Mills, D., Robson, J., & Xu, X. (2025). *Degrees of Difference: inter-generational social mobility and higher educational internationally – a comparative perspective*. The Sutton Trust. Available at: <https://www.suttontrust.com/wp-content/uploads/2025/11/Degrees-of-Difference.pdf>
- Brown, R. (2015). The marketisation of higher education: Issues and ironies. *New Vistas*, 1(1), 4–9.
- Brown, R., & Carasso, H. (2013). *Everything for sale? The marketisation of UK higher education*. Routledge.
- Bunce, L., Baird, A., & Jones, S. E. (2017). The student-as-consumer approach in higher education and its effects on academic performance. *Studies in Higher Education*, 42(11), 1958–1978. <https://doi.org/10.1080/03075079.2015.1127908>
- Chowdry, H., Crawford, C., Dearden, L., Goodman, A., & Vignoles, A. (2013). Widening participation in higher education: Analysis using linked administrative data. *Journal of the Royal Statistical Society. Series A, Statistics in Society*, 176(2), 431–457. <https://doi.org/10.1111/j.1467-985X.2012.01043.x>
- Cuesta-Delgado, D., Barberá-Tomás, D., & Marques, P. (2025). A text-mining analysis of Latin America Universities' mission statements from a 'Third Mission' perspective. *Studies in Higher Education*, 50(7), 1420–1438. <https://doi.org/10.1080/03075079.2024.2377371>
- Daenekindt, S., & Huisman, J. (2020). Mapping the scattered field of research on higher education. A correlated topic model of 17,000 articles, 1991–2018. *Higher Education*, 80, 571–587. <https://doi.org/10.1007/s10734-020-00500-x>
- Dearing, R. (1997). *Higher education in the learning society: Report of the National committee of inquiry into higher education*. Her Majesty's Stationery Office.
- Donnelly, M. (2014). The road to Oxbridge: Schools and elite university choices. *British Journal of Educational Studies*, 62(1), 57–72. <https://doi.org/10.1080/00071005.2014.902028>
- Feinerer, I., Hornik, K., & Meyer, D. (2008). Text mining infrastructure in R. *Journal of Statistical Software*, 25(5), 1–54. <https://doi.org/10.18637/jss.v025.i05>
- Fromm, E. (1978). *To have or to be?* Cape.
- Grün, B., & Hornik, K. (2011). topicmodels: An R package for fitting topic models. *Journal of Statistical Software*, 40(13), 1–30. <https://doi.org/10.18637/jss.v040.i13>
- Harrison, N. (2018). Using the lens of 'possible selves' to explore access to higher education: A new conceptual model for practice, policy, and research. *Social Sciences*, 7(10), Article 209. <https://doi.org/10.3390/socsci7100209>
- Hart, B. (2024). Tremendous pressure: Exploring the pressures faced by widening participation evaluators. *Widening Participation and Lifelong Learning*, 26(2), 230–247. <https://doi.org/10.5456/WPLL.26.2.230>
- House of Lords (2023). *Must do better: The Office for Students and the looming crisis facing higher education* (2nd Report of Session 2022–23, HL Paper 246). House of Lords.
- Jones, S., Hordósy, R., Mittelmeier, J., Quyoum, A., & McCaldin, T. (2021). Possible selves' in practice: How students at further education colleges in England conceptualise university. *Research Papers in Education*, 37(6), 757–772. <https://doi.org/10.1080/02671522.2020.1864765>
- Kandiko Howson, C., Cohen, E., & Viola, J. K. (2022). Inertia in elite STEM widening participation: The use of contextual data in admissions. *British Journal of Sociology of Education*, 43(6), 950–969. <https://doi.org/10.1080/01425692.2022.2085660>

- Lyotard, J. F. (1984). *The postmodern condition: A report on knowledge*. Manchester University.
- Marginson, S. (2011). Higher education and public good. *Higher Education Quarterly*, 65(4), 411–433. <https://doi.org/10.1111/j.1468-2273.2011.00496.x>
- Marginson, S. (2019). Limitations of human capital theory. *Studies in Higher Education*, 44(2), 287–301. <https://doi.org/10.1080/03075079.2017.1359823>
- McCaig, C. (2010). Access agreements, widening participation and market positionality: Enabling student choice? In M. Molesworth, E. Nixon, & R. Scullion (Eds.), *The marketisation of higher education and the student as consumer* (pp. 115–129). Routledge.
- McCaig, C. (2015). The impact of the changing English higher education marketplace on widening participation and fair access: Evidence from a discourse analysis of access agreements. *Widening Participation and Lifelong Learning*, 17(1), 5–22. <https://doi.org/10.5456/WPLL.17.1.5>
- McCaig, C. (2025). Policy Context: The Political Economy of Access and Success in the English Market. In L. Austen & C. McCaig (Eds.), *Delivering and Evaluating Participation after Access : Higher Education in a Marketised System* (pp. 9–33). Emerald.
- McCaig, C., & Adnett, N. (2009). English universities, additional fee income and access agreements: Their impact on widening participation and fair access. *British Journal of Educational Studies*, 57(1), 18–36. <https://doi.org/10.1111/j.1467-8527.2009.00428.x>
- Molesworth, M., Nixon, E., & Scullion, R. (2009). Having, being and higher education: The marketisation of the university and the transformation of the student into consumer. *Teaching in Higher Education*, 14(3), 277–287. <https://doi.org/10.1080/13562510902898841>
- Moore, J., McNeill, J., & Halliday, S. (2011). Worth the price? Some findings from young people on attitudes to increases in university tuition fees. *Widening Participation & Lifelong Learning*. <https://doi.org/10.5456/wpll.13.1.57>
- Naidoo, R. (2004). Fields and institutional strategy: Bourdieu on the relationship between higher education, inequality and society. *British Journal of Sociology of Education*, 25(4), 457–471. <https://doi.org/10.1080/0142569042000236952>
- Oliver, T. J. (2025). Learning to stand on its own two feet: The Office for Students and the Crisis in Higher Education in England. *Political Quarterly*, 96(1), 206–210. <https://doi.org/10.1111/1467-923X.13503>
- Parliament (2023). *Written evidence from Independent Higher Education, IHE (WOS0065)*. <https://committees.parliament.uk/writtenevidence/121147/html/>
- Patrick, F. (2013). Neoliberalism, the knowledge economy, and the learner: Challenging the inevitability of the commodified self as an outcome of education. *ISRN Education*, Article 987521. <https://doi.org/10.1155/2013/108705>
- Pickering, N. (2019). No such thing as a free lunch: Widening participation in English higher education for those on free school meals. *Widening Participation and Lifelong Learning*, 21(3), 57–80. <https://doi.org/10.5456/WPLL.21.3.57>
- QAA (2007). *Guidance for the admissions of students to QAA-recognised Access to HE programmes*. <https://web.archive.org/web/20120818164708/http://www.accesstohe.ac.uk/home/publications/AdmissionGuidanceStudents/default.asp>
- Robinson, D., & Salvestrini, V. (2020). *The impact of interventions for widening access to higher education*. https://epi.org.uk/wp-content/uploads/2020/01/Widening_participation-review_EPI-TASO_2020.pdf
- Teixeira, P., & Dill, D. D. (2011). *Public vices, private virtues? Assessing the effects of marketization in higher education*. Sense.
- Tomlinson, M. (2018). Conceptions of the value of higher education in a measured market. *Higher Education*, 75, 711–727. <https://doi.org/10.1007/s10734-017-0165-6>
- Wakeling, P., & Savage, M. (2015). Entry to elite positions and the stratification of higher education in Britain. *Sociological Review*, 63(2), 290–320. <https://doi.org/10.1111/1467-954X.12284>
- Younger, K., Gascoine, L., Menzies, V., & Torgerson, C. (2019). A systematic review of evidence on the effectiveness of interventions and strategies for widening participation in higher education. *Journal of Further and Higher Education*, 43(6), 742–773. <https://doi.org/10.1080/0309877x.2017.1404558>