

Ofsted inspection reports in early childhood education settings narrowly focussed: A corpus and sentiment analysis

Journal of Early Childhood Research

1–15

© The Author(s) 2025



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1476718X241306859

journals.sagepub.com/home/ecr**Susana Castro-Kemp** 

IOE, UCL's Faculty of Education and Society, UK

Peter Kemp

King's College London, UK

Abstract

High-quality early childhood education is key for children's positive and long-term outcomes. Countries differ on how they quality assure provision. In England, Ofsted is the independent inspection body, often the source of controversy for lack of transparency in ratings. However, there is a dearth of evidence on how inspectors' judgements are made, particularly in the early years sector. This study examines the consistency and nature of individual judgements in Ofsted reports, through an in-depth corpus analysis of inspection reports for non-domestic early years providers (nurseries and preschools) in England. It is the first to analyse the extent to which inspections focus on well-regarded understandings of high-quality early childhood pedagogy and provides an appraisal of the affective tone of the reports, using sentiment analysis. Results show that inspectors are fairly consistent, with similar high frequent terms present across reports, but the inspection focus may be too limited in scope. There is an over-emphasis on school readiness based on outcomes of older children, without clear references to the participation and engagement of younger children and process variables. A more positive tone in the reports is more frequent in high-quality rated settings, except for the section describing children's outcomes, the shortest across reports.

Keywords

early childhood education and care, inspection, Ofsted, quality

Corresponding author:

Susana Castro-Kemp, Psychology and Human Development, IOE, UCL's Faculty of Education and Society, 20 Bedford Way, London WC1E 6BT, UK.

Email: S.Castro-Kemp@ucl.ac.uk

Introduction

The importance of early childhood education is widely recognised. The well-regarded Heckman Curve shows that investments in early years bring long-term societal benefits (Heckman et al., 2013). Despite this, expenditure on Early Childhood Education and Care (ECEC) varies substantially even amongst western European countries with high Gross Domestic Product (GDP; OECD, 2023). For example, pre-primary spending is highest at just under 1.0% of GDP in Iceland, over 0.5% of GDP in Denmark, France, Finland, Iceland, Norway and Sweden, and just under 0.5% in the United Kingdom (UK; Organisation for Economic Cooperation and Development; OECD, 2023). Since 2009–10, spending on free access to ECEC in the UK has more than doubled, as the government increased the number of funded universal hours and brought in new entitlements (Farquharson et al., 2023), bringing to the fore-front the question of what constitutes high-quality provision.

In England, specifically, the Early Years Foundation Stage curriculum (EYFS) extends from birth to 5 years of age, at which point most children attend reception year (which is part of primary schooling, but non-compulsory; Department for Education [DfE], 2021). Before the age of 5, there are few maintained ECEC settings. Recent data from the Department for Education's (DfE) provider survey suggests that 52% of ECEC settings are within the private/independent sector and 12% are childminders (DfE, 2022), with some families having been entitled to government support depending on income (e.g. tax-free childcare or up to 30 hours free childcare per week for 3 and 4-year-olds). In addition to this, the 40% most disadvantaged 2-year-olds are also entitled to a funded part-time place (Farquharson et al., 2023). There are also 14% of places in voluntary-group based providers and 22% which are school-based providers (DfE, 2022). More recently, the government has introduced a gradual increase to ECEC entitlement: by September 2025, most working families with children under the age of 5 will be entitled to 30 hours of childcare support, as well as eligible working parents with a child from 9 months. These changes apply to all UK nations. With more access to ECEC providers, the quality of provision becomes of greater concern.

The quality of all education settings, including early years providers, in England, is monitored by the Office for Standards in Education, Children's Services and Skills (Ofsted). Ofsted carries out periodic inspections of providers, classifying them as Outstanding, Good, Requires Improvement or Inadequate. Ofsted is non-ministerial and independent, but it reports to Parliament. Before Ofsted, inspections were conducted by Local Authorities. With the introduction of the Education (School) Act 1992, and to respond to concerns about variable inspection standards, Ofsted was created and then its remit extended to childcare and childminding in 2001. In 2008, with the introduction of the Early Years Foundation Stage (EYFS), Ofsted inspections started to be carried out against that standard which regulates provision in learning and welfare for any child from birth to 5 years of age.

Over the years, Ofsted has been the source of considerable controversy, including for being over-reliant on numbers rather than on more nuanced aspects of quality, for the fact that many inspectors lacked the knowledge and experience of teaching that some consider essential to perform an accurate inspection, and for causing serious and avoidable stress in the workforce, for not focussing enough on recognised indicators of high-quality ECEC, for conducting outdated evaluations and for putting undue stress on management. Roberts-Holmes (2014), for example, argues that the current early years provision has been narrowed to respond to increased inspection and surveillance, which intensifies pressure to focus on 'school readiness' rather than holistic development; the author also argues that early years pedagogy has become too data-driven and that this disrupts and constrains early years teachers' deeply held child-centred pedagogical values, therefore suggesting that Ofsted inspections may be counter-productive. However, there is little to no

evidence available to say that this is the case, by systematically examining the consistency and quality of judgements made by Ofsted inspectors and their alignment with principles of high-quality ECEC and recognised Early Years pedagogy. Other sources of evidence provide a strong empirical basis for the claim that Ofsted inspectors' judgements are not consistently delivered, for example, the work of Bokhove et al. (2023), which draws on over 30,000 reports, concludes that male inspectors award slightly more lenient judgements to primary schools than their female counterparts, and inspectors employed as permanent staff (Her Majesty's Inspectors) are found to be harsher than those who inspect schools on a freelance basis. Consequently, given its somewhat arbitrary nature, it becomes understandable that school inspection outcomes have few links to student outcomes. von Stumm et al. (2021), have demonstrated that school quality ratings are not good predictors of student achievement, based on a population-based sample (the National Pupil Database) of over 4000 pupil and their schools. Lastly, empirical data is available resulting from the combination of School Census data and Ofsted inspections in England showing that children from more disadvantaged areas have access to better-qualified staff, largely because they are more likely than children from richer areas to attend maintained nursery classes staffed by teachers, and less likely to attend services in the private, voluntary and independent sectors; however, despite this, within both maintained and private/independent sectors, services catering for more disadvantaged children receive lower quality ratings from Ofsted, and a higher concentration of children from disadvantaged areas appears to reduce the likelihood of higher Ofsted grades (Gambaro et al., 2015). The authors suggested that this may be because Ofsted favours discrete child development outcomes (designated as 'school readiness') rather than processual variables related to high-quality provision, and consequently, nurseries with children who start from a vantage point, developmentally, will perform better in Ofsted evaluations. The evidence above refers to primary and secondary school inspections, with little research to date focussing on early years provision. Within early years, Blanden et al. (2017) analysed large administrative datasets and concluded that although children's outcomes are related to the nursery they attend, correlations with Ofsted ratings are weak, suggesting that Ofsted's focus on developmental outcomes does not annul the predictive effect of deprivation versus advantage on children's outcomes.

This body of evidence contributed to several recent changes introduced in the system of Ofsted inspections. The New Inspection Framework was introduced in 2019 and currently regulates how inspections are conducted in all education settings, including early years providers. The main changes introduced with this new framework include deferring an inspection when children are not on roll, the timeline for when inspections may be conducted and their frequency (which is now higher), more emphasis on the settings' capacity to improve, and more emphasis on safeguarding. In the early years, the areas of assessment are now listed as overall effectiveness, the quality of education, behaviour and attitudes, personal development and leadership and management. Inspections of ECEC settings are normally conducted with little notice (often 1 day), and carried out on 1 single day. They may include observations of children's play and interactions with adults, conversations with management and parents/carers and analysis of reports and policies (Ofsted, 2019). The latest changes introduced in the inspectors' handbook refer to the inspectors' conduct, which should be one of encouragement and constructive feedback (Ofsted, 2024). Despite these, the inspection process remains essentially the same: internationally, inspections of educational settings have been described as positioned towards either end of a continuum between 'high stakes sanction oriented' and another end that is 'low stakes advisory focussed' (Simeonova et al., 2020), with Ofsted clearly positioned near the high stakes end, despite regular updates in response to pressure from stakeholders.

The issue of what is considered high-quality ECEC is still not sufficiently described in the Inspectors' handbook or within the new inspection framework, with little evidence of links between

the procedure adopted and the vast body of scientific literature available in the field (e.g. Leach et al., 2008; Mathers et al., 2014; Melhuish et al., 2019; Ulferts et al., 2019), often challenging the concept of ‘school readiness’ as a set of discrete skills, such as emerging literacy and numeracy needed by a certain age to be able to transition to formal primary school; the latter contradicts the understanding of development as unique and individual (Bradbury, 2020). More recent conceptualisations of school readiness encompass a range of outcomes, including academic-related skills and also social and emotional outcomes, and the EYFS definition of Good Level of Development (GLD) includes a range of outcomes across all developmental areas, which are regarded as a measure of school readiness (Kay, 2022). However, a recent Ofsted paper – ‘Bold Beginnings’, commissioned to review the leaders’ understandings of children’s outcomes and their impact on provision, has been critically regarded as reinforcing that the purpose of ECEC as to make children ready for school (Kay, 2022).

One other important point to note in relation to inspections of provision in early years is that, understandings of high-quality early childhood education and care are culturally dependant; in multicultural settings, common in England, expectations of children’s outcomes may and should vary to reflect the wide variety of cultures and attitudes of the children and their families (Ang, 2010).

Although grading descriptors are available to support inspectors in their evaluation of the quality of the settings, handbooks state that inspectors ‘must use their professional judgement to interpret and apply the grade descriptors to the setting’ (Ofsted, 2019, Part I, clause 120).

The purpose of this study was to respond to the dearth of evidence on the nature of individual judgements in Ofsted reports resulting from inspections to early years providers by examining the consistency and nature of individual judgements in Ofsted reports, through an in-depth corpus and sentiment analysis of reports for non-domestic early years providers (nurseries and preschools) in England. The following research questions were formulated: (a) what are the specific domains of focus in the inspection reports for early years providers in England within each of their grading areas? (b) how do the Ofsted reports for early years providers in England align with recognised principles of high-quality pedagogy in early years? (c) What is the nature of the tone used in Ofsted reports and how does this relate to the setting’s rating in each of the grading criteria?

Methods

Document sampling

This study examined all Ofsted inspection reports of non-domestic childcare providers included in the Department for Education (DfE) Childcare providers and inspections list as of 31 March 2022, including all inspections from 2015. This was decided to ensure we would gather the most up to date inspections at the time that the study was conducted; if a setting in the DfE list had an inspection older than 2015 (5 years), it would be excluded. Ofsted inspects settings every 5–7 years or less. Provider reports were taken from across the country, covering 112 local authorities and the work of 197 individual inspectors. Ofsted reports are publicly available and downloadable from the Ofsted’s Website at <https://reports.ofsted.gov.uk/>. The sample in this study included 42 ‘Outstanding’ rated settings (15.1%), 198 ‘Good’ settings (71%) and 39 settings rated with ‘Requires Improvement’ (14%).

Procedure and data analysis

Corpus analysis was conducted using R software. Corpus analysis has been widely used in the field of Linguistics, since the 1960s; it developed from the premise that text or the corpus should be

analysed in its natural environment, with minimal experimental interference (Hunston, 2006). The most common technique within this method is the identification of frequency counts of items encountered in a text, enabling the researcher to spot key words and phrases; another analytic technique involves the examination of words that occur around the most frequently identified ones - their concordances. Lastly, collocation analysis enables the researcher to identify and extract terms within a text that are correlated (or, in other words, that collocate) with any other word, enabling an examination of how words are used in context. These methods are processed quantitatively, through statistical processing of corpus (sets of text) data. However, quantitative results are interpreted in a qualitative fashion by the analyst, based on theory and evidence available (Brezina, 2018).

In this study we conducted frequency and collocation analysis of Ofsted reports, the latter using collocations of 3–5 words, for a better in-depth understanding of the context in which the terms are used, beyond just their frequency. For consistency, we did so within the sections referring to (1) Effectiveness of leadership and management, (2) Quality of teaching, learning and assessment, (3) Personal development, behaviour and welfare and (4) Outcomes for children, the areas under analysis as per Ofsted 2021 guidance, and to ensure we always looked at the same sections across reports. Moreover, the corpus of reports includes other text data not relevant for the purpose of the study. By only including these grading-related sections, we enabled a more fine-grained and focussed analysis.

In addition to corpus analysis, sentiment analysis (known also as opinion mining or emotion AI) was conducted. This looks to systematically identify, extract, quantify and study affective states and subjective information within text. One form of basic sentiment analysis is classifying the polarity of a text regarding whether statements are positive, negative or neutral. More advanced sentiment classification looks, for example, at emotional states such as enjoyment, anger, disgust, sadness, fear and surprise (Pang and Lee, 2008). Sentiment analysis has been used before by Bokhove and Sims (2021) to examine the complete corpus of Ofsted reports released by the English national schools' inspectorate since the turn of the century and concluded that the tone of the reports is more negative in lower-quality rated schools. In this study, we conducted sentiment analysis of Ofsted reports looking specifically at the polarity of the corpus, per Ofsted grading area - (1) Effectiveness of leadership and management, (2) Quality of teaching, learning and assessment, (3) Personal development, behaviour, and welfare and (4) Outcomes for children. We used the Silge and Robinson's (2017) approach to text-mining with R, and specifically the AFINN dataset for sentiment analysis, which is a lexicon of English words rated for valence with an integer between minus five (negative) and plus five (positive), which have been manually labelled by Nielsen (2011). In practice, this is a two-column table where words have been manually matched to a positive or negative sentiment; for example, 'abilities' were matched to 'positive' (score 2), within a scale from -5 to 5.

Results

The purpose of this study was to examine the consistency and nature of individual judgements in Ofsted reports, through an in-depth corpus analysis of inspection reports for non-domestic early years providers (nurseries and preschools) in England. The study examines, for the first time, the extent to which those reports focus on well-regarded understandings of high-quality early childhood pedagogy. Additionally, the study aimed to examine the affective tone of the reports, using sentiment analysis.

The grading area with the highest frequency of words was Quality of teaching, learning and assessment ($n=279$; $M1=127.26$; $SD2=20.7$), followed by Effectiveness of leadership and

management ($n=279$; $M=126.74$; $SD=21.34$), then Personal development, behaviour and welfare ($n=279$; $M=113.68$; $SD=21.34$), and lastly Outcomes for children ($n=279$; $M=87.63$; $SD=82.95$).

Areas of focus in the inspection reports for early years providers in England per grading criteria and Ofsted rating

As a first step we looked at word frequency within each of the grading areas. Table 1 shows the words most frequently adopted in around 60% of the reports per grading area. The most frequent words within Leadership and management, and covered in almost all the reports, are ‘safeguarding’ and ‘effective’, followed by ‘staff’, ‘children’ and ‘children’s’, ‘training’ and ‘procedures’; Here, the most frequent words are the same across Outstanding, Good and Requires Improvement settings.

Within Quality of teaching, learning and assessment (see Table 1), almost all reports refer to ‘children’ and ‘staff’, followed by ‘children’s’, ‘learning’, ‘example’, ‘use’, ‘skills’, ‘activities’ and ‘support’, and this is across Ofsted ratings, with the exception of ‘Requires Improvement’ rated settings which all refer to ‘requires’ and ‘improvement’.

Within Personal development, behaviour and welfare all reports refer to ‘children’, and most reports refer to ‘staff’, followed by ‘children’s’, ‘well’ and ‘example’ (see Table 1); the top frequent words do not differ between Ofsted ratings.

Lastly, within Children’s outcomes the most frequent words are ‘children’, ‘skills’, ‘learning’, ‘school’, ‘make’, ‘progress’ and ‘good’ (see Table 1). This is similar across Ofsted ratings.

As a second step, we looked at collocations – sets of words that appear together frequently, to provide additional information on the context in which they were used. We used 3–5 words to provide as much information on the context as possible. Table 2 shows the frequency of these collocations per grading area. Here, we highlight the frequency of safeguarding related collocations within Leadership and management, which is similar across Ofsted ratings.

Within Quality of teaching, learning and assessment we observe a focus on language and communication, special educational needs and disabilities (SEND), English as an additional language (EAL) and monitoring of children’s learning, and this is present across both ‘Outstanding’ and ‘Good’ settings. For ‘Requires improvement’ settings the terms ‘requires improvement’ have been used most frequently.

Within Personal development, behaviour and welfare we see an emphasis on staff as role models for the children and on children’s self-care skills. Lastly, within Children’s outcomes, the focus is on observable progress and the next phase of formal schooling. Here, the frequency of 5-word collocations denotes a similar focus across Ofsted ratings – school readiness: the most frequent collocations within ‘Outstanding’ settings are ‘are exceptionally well prepared for’ (19%), ‘for the next stage’ (14.3%) and ‘progress from their starting points’ (14.3%). Within ‘Good’ settings, the most frequent collocations are: ‘make good progress from their’ (14.5%), ‘progress from their starting points’ (14%), ‘all children make good progress’ (13%). Lastly, within ‘Requires Improvement’ settings the most frequent collocations are: ‘make the best possible progress’ (16.7%) and it should be noted that ‘children do not make the’ [progress] appears in 10% of reports here, and this Ngram does not appear in other ratings.

It should be noted that the proportions are small when looking for 5-word collocations, because these repeated patterns will only appear in a minority of reports. As we looked for shorter collocations, of 3 words, for example, proportions rise, but the overall context remains the same, thus supporting the interpretation of the meaning of 5-word collocations. For example, when looking at 3-word collocations within the Children’s outcomes grading area, we see that ‘make good

Table 1. Word frequency per grading area.

Word	Leadership & management		Quality of teaching		Personal development, behaviour and welfare		Children's outcomes				
	Frequency	% of reports	Word	Frequency	% of reports	Word	Frequency	% of reports			
safeguarding	377	99.6	children	1330	99.3	children	1511	100	children	1100	99.3
effective	370	99.3	staff	1004	97.8	staff	781	95.7	skills	371	81.4
staff	1201	97.8	children's	586	90.0	children's	316	67.7	learning	322	75.6
children	667	92.5	learning	398	80.6	well	272	65.9	school	220	75.3
children's	331	70.6	example	322	75.6	example	253	60.6	make	255	68.5
training	266	69.9	use	313	67.7	good	255	57.0	progress	215	65.9
procedures	231	62.0	skills	287	65.6	learn	215	54.8	good	284	62.4
manager	311	59.5	activities	249	61.6	help	196	48.4	develop	244	58.8

Table 2. Collocations of 5 words and their frequency per grading area.

Collocations	Leadership and management		Quality of teaching learning and assessment		Personal development, behaviour and welfare		Children's outcomes	
	Count and proportion	Collocations	Count and proportion	Collocations	Count and proportion	Collocations	Count and proportion	Collocations
arrangements for safeguarding are effective	92 (33%)	English as an additional language	20 (7.2%)	staff are good role models	17 (6.1%)	progress from their starting points	40 (14.3%)	
the arrangements for safeguarding are	74 (26.5%)	children's next steps in learning	18 (6.5%)	how to keep themselves safe	11 (3.9%)	make good progress from their	31 (11.1%)	
for safeguarding are effective staff	43 (15.4%)	children's communication and language skills	14 (5%)	learn about the importance of	11 (3.9%)	English as an additional language	28 (10%)	
concerns about a child's welfare	35 (12.5%)	speak English as an additional	13 (4.7%)	children have good opportunities to	10 (3.6%)	all children make good progress	27 (9.7%)	
suitable to work with children	31 (11.1%)	who speak English as an	13 (4.7%)	what is expected of them	10 (3.6%)	good progress from their starting	27 (9.7%)	

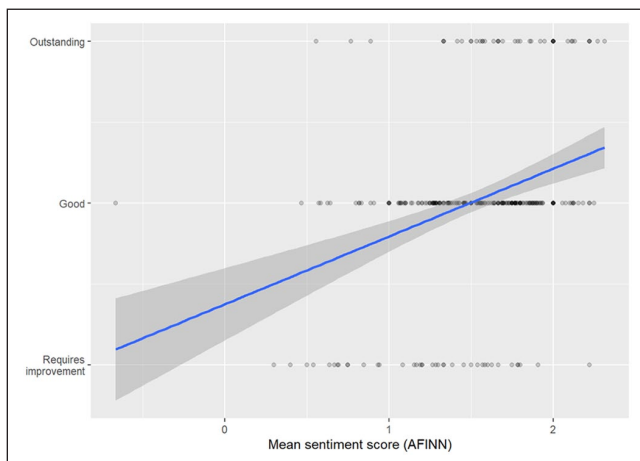


Figure 1. Relationship between sentiment observed in the reports and Ofsted rating, within Effectiveness of leadership and management.

progress' (34.8%), 'in their learning' (25.4%), 'children make good' (21.9%) are the most frequent, thus suggesting the emphasis on the next phase of schooling and school readiness is definitely present, and here in a higher proportion of reports (as we required a shorter Ngram). Similarly, 3-word collocations for Effectiveness of leadership and management support that there is here a focus on effectiveness of safeguarding ('safeguarding is effective' is the most frequent 3-word collocation in 'Outstanding' (66.7%), 'Good' (62.1%), and Requires Improvement settings (84.6%)). The most frequent 3-word collocations within Personal development, behaviour and welfare are also the same across Outstanding and Good settings ('e.g. children', present in 27.3% and 31% of reports respectively), denoting an emphasis on what the staff do to promote children's development and wellbeing, however the most frequent one in Requires Improvement settings is 'staff do not' (40.5%), followed then by 'for example children' (27%).

The affective tone used in Ofsted reports and its relation to the settings' ratings per grading area

The study examined the affective tone of the reports, using sentiment analysis. Figures 1 to 4 show the relationship between Ofsted rating and sentiment in the reports. There is a positive and significant association between Ofsted rating (using Good and Outstanding ratings) and sentiment, across all Ofsted areas of focus, except in relation to Children's outcomes. This means that the more positive the affective tone is, or 'sentiment' of the report, the more likely is the respective setting to be rated Outstanding, except for the area of children's outcomes, where this relationship of positive sentiment associated with higher Ofsted ratings is not observed. While the affective tone is usually more positive in the highest rated providers in relation to Leadership and management ($r(277)=-5.81, p=0.00$), Quality of teaching ($r(277)=-2.87, p=0.00$) and in Personal development, behaviour and welfare ($r(277)=-2.87, p=0.00$), such relationship is not observable in the Children's outcomes area, where the sentiment is always the same, across Ofsted ratings ($r(276)=-0.41, p=0.09$).

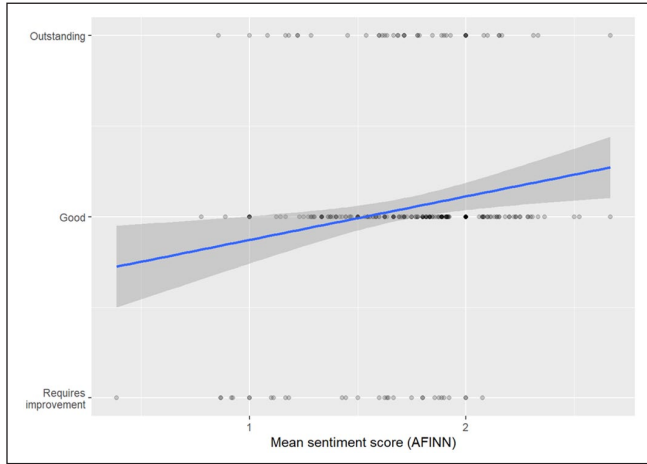


Figure 2. Relationship between sentiment observed in the reports and Ofsted rating, within Quality of teaching, learning and assessment.

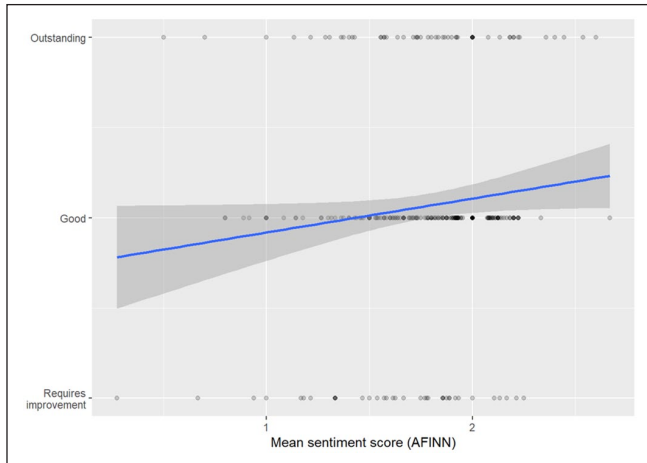


Figure 3. Relationship between sentiment observed in the reports and Ofsted rating, within Personal development, behaviour and welfare.

Discussion

The purpose of this study was to address the lack of research evidence on the consistency and focus of individual judgements in Ofsted reports, through an in-depth corpus and sentiment analysis of inspection reports for non-domestic early years providers (nurseries and preschools) in England.

Results from the frequency analysis and collocation analysis conducted showed that within each grading area, it is possible to identify clear specific concerns emphasised by Ofsted. For example, when looking at the Leadership and management area, inspectors seem to focus mainly on safeguarding issues and whether staff are made confident and skilled to spot and act when necessary. There are no references to strategy to maintain staff and avoid high turnover, which is a recommended practice and a current pressing issue in ECEC in England (Ryan et al., 2017; Sylva and

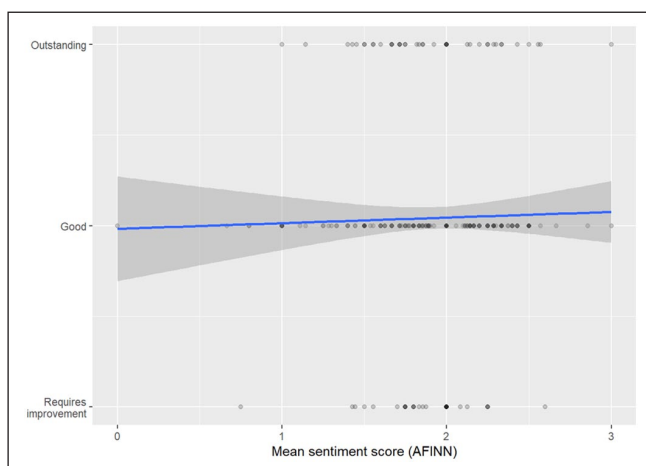


Figure 4. Relationship between sentiment observed in the reports and Ofsted rating, within Children's outcomes.

Pugh, 2005). The area of Quality of teaching, learning and assessment shows an overall focus on language and communication, special educational needs and on how staff monitor children's learning. Within Personal development, behaviour and welfare, Ofsted focuses mainly on the extent to which staff act as a role model of children's behaviour, encouraging their independence. Lastly, the Children's outcomes area clearly shows an emphasis on school readiness understood as progress made towards the next phase of formal education, and while this is the shortest area across reports, it is also the one where a specific focus is very apparent. This suggests that Ofsted's understanding of children's outcomes is mostly concerned with academic outcomes needed for formal schooling. This interpretation of the context of word frequency was supported by reducing the number of collocated words to 3, which showed a similar frequency of content across areas, with higher frequency proportions (as patterns of 3 words are naturally more frequently occurring than patterns of 5). We also observed that this content inclination within each area is independent of the settings' Ofsted rating, apart from those with 'Requires Improvement' ratings, where collocations denote an emphasis on actions that staff do not undertake, and presumably should be undertaking.

These results suggest, on the one hand, that there is a substantial amount of consistency between Ofsted inspectors in what constitutes the focus of their observations. Clearly the training received, and materials provided to inspectors do enable a considerable level of homogenous thinking when performing judgements. However, on the other hand, the fact that consistency is observed is not a synonym of inspections aligned with recognised early years pedagogy. The Early Years Inspectors' Handbook (e.g. Ofsted, 2019, 2023) provides considerable detail on how to make judgements and use the rating scale, however, it does not go as far as to which specific areas of everyday life in the setting one should observe. For example, while the handbook is clear that inspectors should focus on 'the progress children make in their learning and development relative to their starting points, and their readiness for the next stage of their education' (Ofsted, 2019, 2023, Part I, point 37), and this is clearly a main focus of the reports issued, we did not observe the same in relation to other Ofsted guidance in the handbook; for example inspectors should focus on children's personal and emotional development, including whether they feel safe and are secure and happy, according to the handbook, but the inspections carried out do not seem to consider evidence-based methods of accessing children's views, and/or this is not reported.

We also know from extensive literature in the field of ECEC that most learning in early years occurs through play (Vogt et al., 2020), and that incidental learning forms a significant part of holistic learning and development in early years (Albritton and Johnson, 2023). This is in direct opposition to approaches to learning in Early Years which are overly focussed on school readiness as academic achievement, and have been widely criticised for underestimating the potential of children's creativity in fostering learning and development that will be useful in life and in school (Bradbury and Roberts-Holmes, 2017). Despite these widely recognised assumptions, our study shows that the most frequently assessed outcomes for children are those related to school readiness conceived as preparation for formal schooling. The most frequent terms, across grading areas, do not include 'play', or 'creativity' nor other terms related to engagement with nature and collaborations with parents, all widely recognised principles of high-quality provision throughout the first few years of life (Mathers et al., 2014). This denotes an understanding of high-quality provision, from the point of view of Ofsted, which is inconsistent with the evidence and long-standing theory available and overly focussed on the transition of older children to formal schooling.

Lastly, the sentiment analysis conducted showed that, as expected based on previous studies (e.g. Bokhove and Sims, 2021), Ofsted rating is associated with the tone adopted in the report, but this is not applicable to the children's outcomes grading area, which is unexpected. One possible explanation for this is that as this area is so very much focussed on school readiness, understood as being prepared for formal academic schooling, and this is such an apparent concern across reports, the tone remains unchanged with ratings; It is also the shortest area in the reports, as attested by word frequency, and so it sometimes only states the level of school readiness, without much more information of qualitative nature specifically on the participation and engagement of the younger children. Therefore, while for the areas concerning leadership and managing, quality of teaching and learning and personal development and welfare the more positive the affective tone of the report is, the more likely that setting is to be rated Outstanding, the Children's Outcomes section does not show this trend. Rather, it focuses solely on children being academically ready for school (or not), without reference to other more nuanced outcomes – for example, whether children engage in most daily activities functionally and positively, and progress against their own developmental trajectories. This suggests a quality evaluation process that dismisses holistic outcomes and is focussed almost solely on academic achievement, even when very young children are being considered. Even when considering school readiness as how well-prepared children are to transition to primary school in terms of academic subjects' mastery, research evidence on school readiness has shown that younger children and babies displaying and developing higher levels of social and emotional wellbeing, for example, are more prepared for formal schooling at the end of preschool (Nakamichi et al., 2021), and therefore the focus of high-quality ECEC should be in providing a secure, socio-emotionally positive climate and opportunities for development. Outstanding settings should be able to demonstrate that. However, the current study shows that Ofsted inspections are not concerned with that type of developmental outcomes, within the children's outcomes section, and also that qualitative outcomes are potentially not relevant to the final judgement on ratings for settings made by Ofsted. While it may be that England's ECEC settings are providing opportunities for holistic development, this is not what Ofsted seemed to be looking at the time of the study. By only focussing on outcomes for the older children and not on processual variables of early learning and development for the younger children (e.g. functioning in everyday life and play), Ofsted is favouring those children from economically affluent areas, where it has been shown children start at a vantage point (Gambaro et al., 2015).

One of the possible limitations of this study is that there have been different versions of the Early Years Inspection Handbook issued by Ofsted in the period covered by the reports under

analysis in this study. However, the changes made in those handbooks are not significant for the purpose of the analysis conducted, and could not affect our results or interpretations. For example, the names of the grading areas have changed, but not their general focus. More significant changes in the Ofsted framework have been introduced in 2023, including, for example, the possibility of follow-up visits with managers, and recommendations on the conduct of inspectors were issued; these are significant changes which probably warrant the need for future analysis of reports, to ascertain any changes in focus and sentiment in the final reports. Lastly, the section on outcomes for children is no longer an area of focus on final reports. Future research should look at the most recent reports to gauge whether the over-reliance on school readiness may still be present under different headings. We also know that between 2019 and 2023 the proportion of Outstanding ratings in ECEC has decreased from 20% to 13% of registered providers. The proportion of Good-rated providers has increased from 76% to 82%. Requires improvement-rated settings decreased from 3% to 2% and Inadequate settings remained stable at 1% (Farquharson et al., 2023). These changes reflect recently introduced changes to the way in which Ofsted operates, including the fact that Outstanding-rated settings are now being inspected more frequently than before, but they do not reflect a change in the focus or sentiment of the analysis, which was this study's aim. More importantly, the model of inspection has not shifted from a high stakes sanction oriented (Simeonova et al., 2020) approach, and so the relevance of our findings stands, despite those changes. The findings become even more timely when considering recent changes to access to free childcare implemented in 2023, which will directly impact the quality assurance procedures of providers, because of increased demand.

Particular strengths are that this is the first study to provide systematic corpus analysis of Ofsted reports of ECEC settings in England, delivering much needed insight on the consistency and nature of Ofsted inspectors' judgements. The methodology is based on replicable coding which can be used in future analysis of future reports issued by Ofsted to ascertain change and examine how changes in the Ofsted may reflect changes of focus in judgements. One further area in need of research in the future is the analysis of Ofsted reports following the EYFS in primary schools, which has not been covered in this particular study.

Conclusion

The purpose of this study was to respond to the dearth of evidence on the consistency and nature of individual judgements in Ofsted reports, through an in-depth corpus and sentiment analysis of inspection reports for non-domestic early years providers (nurseries and preschools) in England. Results showed that although Ofsted reports appear fairly consistent, with high frequent terms present in large proportions of the reports, the focus of the inspections may be limited. Data shows that the inspectors' comments do not align with the evidence base on what constitutes high-quality early childhood education and care: there is an over-emphasis on school readiness as academic preparedness for the next phase of formal schooling, but no evidence of preoccupation with participation and engagement of the younger children and processual variables of ECEC throughout the whole of Early Years education as specific outcomes. This probably reflects an inspection process that is limited in time and scope, not allowing for deep engagement with stakeholders, which would favour a much more detailed evaluation of the criteria at stake.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Froebel Trust under Grant number RCH-AP-00154-2019.

ORCID iD

Susana Castro-Kemp  <https://orcid.org/0000-0002-9838-7316>

Data availability statement

The code used to retrieve and perform the analysis of the Ofsted reports in R software is available upon reasonable request to the corresponding author.

References

- Albritton K and Johnson L (2023) Onset-rime, alliteration, and incidental teaching. In: Coolong-Chaffin M, Hawkins R and Axelrod M (eds) *Reading Intervention Case Studies for School Psychologists: Evidence-Based Implementation and Analysis*, 1st edn. New York: Routledge, pp.20–39.
- Ang L (2010) Critical perspectives on cultural diversity in early childhood: Building an inclusive curriculum and provision. *Early Years* 30(1): 41–52.
- Blanden J, Hansen K and McNally S (2017) *Quality in early years settings and children's school achievement*. CEP Discussion Paper No. 1468. London: Centre for Economic Performance. ISSN 2042-2695.
- Bokhove C, Jerrim J and Sims S (2023) Are some school inspectors more lenient than others? *School Effectiveness and School Improvement* 34(4): 419–441.
- Bokhove C and Sims S (2021) Demonstrating the potential of text mining for analyzing school inspection reports: A sentiment analysis of 17,000 Ofsted documents. *International Journal of Research & Method in Education* 44(4): 433–445.
- Bradbury A (2020) Datafied at four: The role of data in the 'schoolification' of early childhood education in England. In: Jarke J and Breiter A (eds) *The Datafication of Education*. Milton Park: Routledge, pp.8–22.
- Bradbury A and Roberts-Holmes G (2017) *The Datafication of Primary and Early Years Education: Playing With Numbers*. New York: Routledge.
- Brezina V (2018) *Statistics in Corpus Linguistics: A Practical Guide*. Cambridge: Cambridge University Press.
- Department for Education (2021) Statutory framework for the early years foundation stage Setting the standards for learning, development and care for children from birth to five. Available at: <https://www.gov.uk/government/publications/early-years-foundation-stage-framework-2>.
- Department for Education (2022) Childcare and early years provider survey. Available at: <https://explore-education-statistics.service.gov.uk/find-statistics/childcare-and-early-years-provider-survey/2022>.
- Farquharson C, Joyce R and Water T (2023) Early years and childcare in England: Public spending, private costs, and the challenges ahead. Institute of Fiscal Studies. Available at: <https://ifs.org.uk/articles/early-years-and-childcare-england-public-spending-private-costs-and-challenges-ahead>.
- Gambaro L, Stewart K and Waldfogel J (2015) A question of quality: Do children from disadvantaged backgrounds receive lower quality early childhood education and care? *British Educational Research Journal* 41(4): 553–574.
- Heckman J, Pinto R and Savelyev P (2013) Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review* 103(6): 2052–2086.
- Hunston S (2006) Corpus linguistics. In: Brown K (ed.) *Encyclopaedia of Language & Linguistics*, 1st edn. Oxford: Elsevier, pp.234–248.
- Kay L (2022) 'What works' and for whom? Bold beginnings and the construction of the school ready child. *Journal of Early Childhood Research* 20(2): 172–184.
- Leach P, Barnes J, Malmberg L, et al.; the FCCC team (2008) The quality of different types of child care at 10 and 18 months: A comparison between types and factors related to quality. *Early Child Development and Care* 178(2): 177–209.
- Mathers S, Eisenstadt N, Sylva K, et al. (2014) Sound foundations: a review of the research evidence on quality of early childhood education and care for children under three. Implications for Policy and Practice. Sutton Trust. Available at: <https://www.suttontrust.com/wp-content/uploads/2019/12/1sound-foundations-jan2014-3-1.pdf>.

- Melhuish E, Barnes J, Gardiner J, et al. (2019) A study of the long-term influence of early childhood education and care on the risk for developing special educational needs. *Exceptionality Education International* 29(3): 22–41. Available at: <https://ora.ox.ac.uk/objects/uuid:f8e88e96-2a29-4269-a576-8cf59ecae58d>.
- Nakamichi K, Nakamichi N and Nakazawa J (2021) Preschool social-emotional competencies predict school adjustment in grade 1. *Early Child Development and Care* 191(2): 159–172.
- Nielsen F (2011) A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. In: *Proceedings of the ESWC2011 workshop on 'making sense of microposts': Big things come in small packages*, pp.93–98. DOI: <https://doi.org/10.48550/arXiv.1103.2903>.
- OECD (2023) PF3.1: Public spending on childcare and early education. *OECD Family Database*. Available at: https://www.oecd.org/els/soc/PF3_1_Public_spending_on_childcare_and_early_education.pdf.
- Ofsted (2019) Early years inspection handbook for Ofsted-registered provision. Available at: <https://www.gov.uk/government/publications/early-years-inspection-handbook-eif/early-years-inspection-handbook-for-ofsted-registered-provision-for-september-2022>.
- Ofsted (2023) Early years inspection handbook for Ofsted-registered provision. Available at: <https://www.gov.uk/government/publications/early-years-inspection-handbook-eif/early-years-inspection-handbook-for-ofsted-registered-provision-for-september-2022>.
- Ofsted (2024) *Early Years Inspection Handbook*. Available at: <https://www.gov.uk/government/publications/early-years-inspection-handbook-eif/early-years-inspection-handbook-for-ofsted-registered-provision-for-september-2023>.
- Pang B and Lee L (2008) *Subjectivity Detection and Opinion Identification*. Hanover, MA: Now Publishers Inc. Available at: <http://www.cs.cornell.edu/home/llee/opinion-mining-sentiment-analysis-survey.html>.
- Roberts-Holmes G (2014) The ‘datafication’ of early years pedagogy: ‘If the teaching is good, the data should be good and if there’s bad teaching, there is bad data.’ *Journal of Education Policy* 30(3): 302–315. <https://doi.org/10.1080/02680939.2014.924561>
- Ryan SV, von der Embse NP, Pendergast LL, et al. (2017) Leaving the teaching profession: The role of teacher stress and educational accountability policies on turnover intent. *Teaching and Teacher Education* 66: 1–11.
- Silge J and Robinson D (2017) *Text Mining With R: A Tidy Approach*. Boston: O’Reilly Media, Inc. Available at: <https://www.tidytextmining.com/index.html>.
- Simeonova R, Parvanova Y, Brown M, et al. (2020) A continuum of approaches to school inspections: cases from Europe. *Pedagogy Critical Approaches to Teaching Literature Language Composition and Culture* 92(4): 487–506. Available at: <https://www.ceeol.com/search/viewpdf?id=881986>.
- Sylva K and Pugh G (2005) Transforming the early years in England. *Oxford Review of Education* 31(1): 11–27.
- Ulferts H, Wolf KM and Anders Y (2019) Impact of process quality in early childhood education and care on academic outcomes: Longitudinal meta-analysis. *Child Development* 90(5): 1474–1489.
- Vogt F, Hauser B, Stebler R, et al. (2020) Learning through play–pedagogy and learning outcomes in early childhood mathematics. In: Thiel O and Perry B (eds) *Innovative Approaches in Early Childhood Mathematics*. London: Routledge, pp.127–141, ebook.
- von Stumm S, Smith-Woolley E, Cheesman R, et al. (2021) School quality ratings are weak predictors of students’ achievement and well-being. *Journal of Child Psychology and Psychiatry* 62(3): 339–348.