


Article

Assessing English: A Comparison between Canada and England's Assessment Procedures

Bethan Marshall * and Simon Gibbons 

King's College London, Department of Education, Communication and Society, Franklin Wilkins Building
Annexe, Waterloo Road, London SE1 9NH, UK; simon.gibbons@kcl.ac.uk

* Correspondence: bethan.marshall@kcl.ac.uk

Received: 5 October 2018; Accepted: 13 November 2018; Published: 5 December 2018



Abstract: English as a subject used to be assessed using course-based or portfolio assessments but now it is increasingly examined through terminal tests. Canada is an exception to this rule. This paper compares the way English is assessed in England and Canada and looks to the ways in which the kind of assessment undertaken affects the practices of English teachers both in the teaching of summative and formative assessment.

Keywords: course-based assessment; summative assessment; formative assessment

1. Introduction

The purpose of this paper is to consider how formative and summative classroom practices varies according to the way terminal, summative assessment is viewed and undertaken by high school teachers in England and Canada. Assessing pupils in English for terminal, high stakes qualifications through the coursework they completed in school was, up until the 1990s, common practice in English speaking countries. Slowly but surely, however, the pattern has changed. England abandoned 100% coursework in 1992, the last pupils to take a 100% summative, course-based exam was in 1994. It maintained some form of coursework until 2015 when any form of course-based exam was eliminated and since 2017, students have been assessed by terminal examination alone. New Zealand toyed with summative, 100% course-based assessment somewhat longer. The University of Waikato ran an English course known as The English Study Design, but this too finished in 2004. Queensland, in Australia, persevered further. They had a system very like the one in England where it combined internal school assessments with external moderation. They even had *Random Sampling of Assessment in Authority Subjects* to double check that the assessment system was robust. This too, stopped recently due, in part, to the Masters Report in 2009, which was actually looking at the National Assessment Plan Literacy and Numeracy (NAPLAN) and Queensland's poor results in that test. The reasoning was that a terminal assessment, which relied on the judgement of teachers, was difficult to maintain. As Klenowski and Wyatt-Smith pointed out,

The traditional divide between objective and subjective judgement became established, the former routinely associated with standardised testing, and the latter, teacher judgement. Underpinning the divide was the ill-conceived notion that standardised testing led to more reliable judgement, especially where marking was regulated (e.g., by machine marking), and relied less on the human brain for decision-making [1] (p. 68).

This divide between objective and subjective assessment lay behind the more cursory attempts at 100% summative, coursework, or portfolio assessment in the United States. Despite some keen advocates (see for example Mellon [2], cited in Freedman [3]; Le Mahieu et al. [4] and Koretz [5]) ultimately, according to Wiliam,

The standards of reliability that had been set by the SATs simply could not be matched with portfolios. While advocates might claim that the latter were more valid measurements of learning, the fact that the same portfolio would get different scores according to who did the scoring made their use for summative purposes impossible in the US context' [6] (p. 178).

Yet despite the abandonment of coursework in England and in other English speaking countries, Canada has maintained a system of course-based assessment for terminal, high stakes exams. Again, it is the purpose of this paper to compare English teachers in Canada and England to explore their experiences of the assessment system they are teaching under and the ways in which they think it affects what they do.

This comparison has particular pertinence in that the most recent Programme for International Student Assessment (PISA), which came out in 2016, placed Canada's score in reading 19 countries ahead of the UK. Canada was second to the UK's 21st. There are many studies that demonstrate that the way students are assessed in PISA is problematic (for example Dohn [7]; Bulle [8]; Pereyra et al. [9], and Murphy [10]; the ways in which countries enter students varies [11,12], and the way in which it impacts on education generally across the world as negative [13]. Our study does not investigate the complexities of the survey, though we do acknowledge the numerous problems with gaining information about teenage reading, most recently in a computerised test. Having said that, in analysing the 2016 results, in England, much was made of Singapore, who came top in reading, but little was said of Canada, which came second. Comparing country's performance is a tricky business. Dylan Wiliam [14] has written a whole chapter in his most recent book entitled 'Copying other countries' in which he highlights the dangers of looking to other countries to find solutions to one's own mediocre performance (in this case he was looking to the States as a point of comparison with other countries). In the end, he claims that, 'looking at what other countries are doing has its place, but the solutions will have to be home-grown' [14] (p. 90).

2. A Brief History

Our comparison between the two countries then, is not one where we say that England should copy the Canadian experience, rather we are simply looking at the ways in which English is assessed as a subject because it is interesting, may be informative, and gives rise to different concerns and practices. The case studies, which we will be looking at, are part of a wider research project, which considers the teaching of high school English in Canada, England, and Scotland but in this article we are looking solely at the assessment practices, both formative and summative, of the first two countries, as their respective contexts and histories differ the most. We collected data from four schools in England and Scotland and three in Canada. We looked at only one province in Canada, Ontario, which has a population of just under 13 million and is the largest province in Canada with the greatest ethnic diversity. Selecting one province, Ontario, to consider may be unrepresentative of Canada as a whole, the provincial government in Alberta, for instance being more regulated than in Ontario, but it does provide a snapshot of a very different way of thinking about how the English curriculum is taught and assessed when compared with the approach in England.

England has a history of government initiatives, including the introduction of a national curriculum, which has been revised five times since it was introduced in 1989; a national literacy hour introduced in 1997 and dropped in 2008 and changes to the examination of its pupils, culminating in the current changes to GCSE which, as we have seen, now omit any form of summative coursework undertaken by the students at all. It is all assessed by terminal examination at the end of a two-year course. Although these changes were not initially designed to improve the PISA ranking, this has latterly been the case [15,16].

Ontario, however, has a curriculum, which is only assessed by the school and the teachers therein. In so doing, as with other Canadian provinces, it is alone amongst the major English speaking jurisdictions. As we have seen, all other countries have some formal system of assessment through which students, and often their schools, are judged. Young and Boyd point out that 'Canada has

not followed other jurisdictions such as the UK and many US States down the road of “competitive certification” [17] (p. 3). Nevertheless, Ontario also introduced a curriculum in 1995, which has undergone minor changes and a literacy test for fifteen-year-olds, which students have to pass in order to graduate from high school. Even the mandatory passing of this test is eased, however, in that you can retake in the year following and if you are still not successful you can take a class in your final year to meet the standards. It is also true, as we shall see, that the test for literacy is seen as very different from the English curriculum. Canada does not have the same level of standards-based reform as any of the other major English speaking countries, particularly England. The English curriculum emphasises reading widely, including texts of the students’ own choice and, throughout, it encourages teachers to make ‘professional decisions’ [18] (p. 18) about all aspects of the curriculum.

3. Research Methodology

We took a qualitative approach to examine our chosen case study schools, including semi-structured interviews [19–21] and lesson observations. These two elements formed the basis of a case study [22–25], although, in this paper, we focus predominantly, but not exclusively, on the interviews with teachers. The interviews were all approximately 45 min, taking place after we had observed a lesson and we asked a range of questions about English teaching including the following: Talk me through the lesson? Questions arising out of the particular nature of the lesson; Has your department been at all influenced by government policy? Does your department have organized schemes of work? Has your department altered its schemes of work and, if so, why? Although none of the above were about assessment, and designed to look at English teaching in general, the topic of assessment arose throughout. The lessons, for example, were typically preparing students for summative exams or for course-based assignments, which would eventually be summatively assessed and so the discussion of the lesson contained much that was insightful about the way they considered student assessment. In this way, we considered how the nature of summative assessment influenced teachers thinking about what they do in the high school English classroom.

Of the three case study schools in Canada, two were in a small university town to the west of Toronto, Cathedral, and Duke, the other, Millennium, was so called ‘blue collar’, and stood in a town to the south of Toronto. The teachers interviewed were respectively Helen, Wanda, and Marian. The schools in England were Churchfields, which was an inner city comprehensive; Woodlands, which was a school at the edge of a large city, both had an ethnically diverse intake; Broadly, which served a suburban, ‘commuter belt’ population in the home counties, and Oakfield which was a village college with a predominantly white middle-class intake. We interviewed four teachers in these schools: Lisa, Catherine, Paul, and Nigel. In addition, we interviewed a Canadian teacher, Chris, actually from Saskatchewan, who was at the time teaching in Oakfield. Case studies do not provide generalizability but allow for a snapshot of practice to be seen from which comparisons and overlaps can be drawn and where a particular case can provide a tentative look at a generalizable principle [22–25].

Within these case-studies, we used arts-based practical criticism using Dewey’s concept of criticism and Eisner’s notions of connoisseurship and critical appraisal. While this may seem subjective, this approach, chosen to align with the long-term practices of the researchers participating in the project, has been validated in two articles, which show how we considered, critiqued, and close read what we observed [26,27]; it has been successfully used in analysing classroom data in an ESRC project [28] and for our book [29]. There is some evidence, also, that English teachers respond better to research when they feel that research is carried out as an arts, as opposed to a traditional social science, approach [30].

Again while acknowledging that some may view criticism as being ‘subjective’ Eisner argues, ‘Each of these concepts, educational connoisseurship, and educational criticism, have their roots in the arts’ [31] (p. 41). Criticism, as Dewey pointed out in *Art as Experience* [32], has at its end, the re-education of perception. What the critic strives for is to articulate or render those ineffable

qualities constituting art in the language that makes them vivid [31] (p. 41). And, he concludes, 'The task of the critic is to help us see' [31] (p. 41).

It is important then, if we are to look at English, to use some of those skills in English in the task of analysing interviews and in studying classrooms—to view both as texts, taking them both apart in a close reading of what was said and done. Thus the art of criticism becomes twofold—it is both an arts discipline and it is about the assessment of texts. Eisner argues that, in part, the job of the critic is to help us see an issue with more clarity than before but also identifies an active role for the reader who should engage with the critique, reflecting on where the reader/viewer stands in relation to it. Using research methodologies that are consistent with the culture of the English classroom is a crucial part of building a culture likely to encourage authentic data. What follows is a critique that comes predominantly from the interviews and, on occasion, the lesson observations, which act as a corroboration of the teachers' interviews, perspectives, and responses.

4. Findings

4.1. *The Differences in the Assessment Systems*

4.1.1. England

If we look at the ways English is summatively assessed in Ontario, Canada, and England, the greatest difference is that those in Ontario assess everything through course-based work while in England pupils are now assessed solely through exams that they take at the end of a two-year course. This means that the interviews of the teachers in England are dominated by talk of the examination system. From a purely objective perspective, one could argue that the national curriculum in England, and in consequence the GCSE specifications, are relatively content-light. Certainly, the explicit rhetoric of the policymakers was that the new national curriculum, introduced in 2013 [33] would specify only the core content and that schools would be free to devise their own broad and balanced curriculum around it, incorporating local interests and concerns as they did this. The whole curriculum, according to the National Curriculum expert panel, should be made up of the national curriculum, the basic curriculum (things like Religious Education), and the local curriculum [34]. Given the freedom from summative coursework, or controlled conditions assessment, and further given that many schools in England will start the GCSE course a year early, it might be suggested that there is plenty of space and time for English teachers so that the seemingly constant preparation for terminal examinations is unnecessary.

The reality is very different; so high are the stakes placed on the GCSE assessments, that reducing content has done nothing to alleviate the relentless focus on preparation for exam type responses. So, for example, they have to grapple with how pupils write in exams: 'I mean it's going to have to become exam heavy because we're going to have to really make sure they've got those exam skills and they feel really comfortable' (Lisa). Her dilemma is that,

I don't think it allows students the time to really engage with a text in a meaningful way and it does mean that often the teaching of English becomes reduced to doing PEE [point, evidence, explanation] paragraphs or EPI [evidence, point, interpretation] paragraphs or PEEL [point, evidence, explanation, link] paragraphs or whatever. I mean there are so many types of these now and students just having to find quotes and kind of you know find a quote and it becomes like that as you get nearer to the exam (Lisa).

It becomes a problem of not what you write but how you write it. The lessons move away from 'engag[ing] with a text in a meaningful way' and become 'reduced to doing PEE paragraphs or EPI paragraphs or PEEL paragraphs or whatever.' The idea of the subject being 'reduced' by the examination process is also significant. Not only is the teacher's job and that of the pupils, constrained, but the subject itself.

The content of what they have to write becomes a concern to English teachers in England too. In Assessment Objective 2 (AO2), pupils have to comment on the form and language in texts. The standard mantra amongst English teachers, for many years, has been that there is little point simply feature spotting—i.e., merely identifying the use of a particular technique or language feature. The value only comes from saying why the writer has used a particular device, or from focusing on the effect of the device on the reader or audience. Yet, much of Nigel's lesson, where he was teaching Steinbeck's *Of Mice and Men* [35], was concerned with pupils determining the linguistic techniques of the writer. This in itself is significant, as he was not teaching an exam class. Traditionally pupils start their GCSE course in year 10, the year above the class being observed but this is an indication that the teachers start preparing them for the demands of the GCSE well before they start the examination course. In fact, it is becoming increasingly common for pupils to start the course a year early, in year 9, so that they have three years to complete the GCSE rather than just two. There is even evidence from inspections [36] that the writing of structured PEE paragraphs begins at the very start of a pupil's secondary school English experience when they are eleven.

Nigel was not, in fact, starting the GCSE course early and yet the importance of knowing the metalanguage, the AO2 element of the GCSE, was clear in both his lesson and interview. In the lesson, he used the analogy of an artist's pallet to talk about the various devices a writer might draw on in the construction of the text, but he was explicit in reminding students of terms within the sphere of diction and imagery that they might draw on. Nigel's attitude to the focus on language was interesting in that he saw the merits of it, and he claimed, to some extent, to be able to be ironic to a degree with his students about the demands. Within the lesson itself there was an exchange with a student, for example, when somewhat self-deprecatingly, but significantly, Nigel confessed the AO2 focus was not something expected of him when he was at school himself. Despite this, however, Nigel's interview showed the perceived importance of the ability to name parts in and of itself and the vocabulary attached is a necessity:

There are things like diction, naming the parts of speech when you feel hearts sinking and think why are we doing this? . . . Where I think it's partly bollocks is doing it with a bunch of 14-year-olds. How do we keep them jogging along when they are going to have to use these things to get a level 5 . . . you have to use the metalanguage? And when I trained I was rather dismayed by a couple of members of staff for whom a whole lesson could be devoted to auditing metalanguage. And I thought god is this it? You know, knowing everything there is to know about bread except you're meant to eat it. So I hope . . . I suppose I do find myself working hard to keep both sides going. The stuff that racks up the marks and the point of it (Nigel).

Nigel explicitly declares that tension—the need for students to be able to utilise metalanguage in order to gain access to the higher echelons of the marking criteria, whilst trying to maintain a sense that response to literature ought to, at its heart, involve some kind of comments on meaning.

The tension is there for Catherine too. The emphasis on commenting on language and form was a governing principle in the way she helped students articulate their responses to the text. In modelling her own paragraph to the class, in order to scaffold their own attempts at writing, the language analysis was highlighted and there was a sense from the teacher that—in the eyes of the examiners at least—the ability to spot features of the language itself is something that would be rewarded. So that, for Catherine, even though some of her students:

Might not have been as good at explaining what they mean, they know what's a simile, they know what's a metaphor, they know repetition . . . they even know something like sibilance but they're just not as good at explaining them. And, to be honest, even though I want them to talk about the plot, a lot of the marks are based on that AO2 . . . that language, form, and structure so I think even if they can recall some of that then that will be useful (Catherine).

It is perhaps a telling comment on the new regime that, in Catherine's view at least, there is actually a reason to merely identify writers' techniques. The result can be to restrain both the ways in which students read texts and the ways in which they feel they are permitted to respond to them.

4.1.2. Canada

Assessments in Ontario is very different. As part of our case studies, we interviewed a Canadian teacher, actually from Saskatchewan, who was at the time teaching in Oakfield. As in Ontario, she explained, that in Saskatchewan, 'We don't have year-end exams set by an exam board, we would do our own. It's all coursework based' (Chris). The only exception to this is the OSSLT (Ontario Secondary Schools Literacy Test), which the English teachers we interviewed did not view as a test of English but more broadly one of literacy. In English, there are no principal assessors examining students' work other than the teachers themselves. Helen, talking of the OSSLT, for example, says 'It's about literacy not about English or literature' (Helen). The difference between it being a test of literacy rather than English is telling because as she points out, 'It's cross-curriculum (ibid).' The result is that, 'We don't teach to the test constantly' Helen). Everything in English, however, is assessed by the teachers within the school and probably by the class teacher alone, with no external check. As Wanda puts it, 'Luckily I mean we don't have standardized tests' (Wanda). When interviewing the teachers, unlike their English counterparts, they do not constantly refer to the examination system. They only do so for clarification with the interviewer. What they do talk about far more are curricular concerns, the way they organise what they teach. Marian, in Millennium, spoke of the way in which English teaching was 'very book focussed'. In Duke, Wanda explained,

Some classes don't even do a novel study like a whole class novel study, they do something more like reading circles or book groups so that students have more choice. So they might get a list of eight titles and they select one of those to be their novel study and then there's three other people in the room reading the same book. And so that can be what the novel study looks like (Wanda).

What is important is that all the schools try and ensure that the students engaged with the books they were studying. Chris emphasises the way that, for her, teaching a coursework based exam, makes teaching pupil orientated: 'You can customise it for your classes', and flexible, 'The rest comes from a combination of attendance and effort or you can do other tests if you like' (Chris). The assessment system also allows her to 'Look into big questions and big ideas as oppose to just building analytical skills' (Chris), typical of the system she finds in England.

The Ontario curriculum for English does, however, prescribe that all high school students have to study Canadian literature. Although there is then a degree of constraint, they have to study Canadian literature, the teachers can select books to which they think the students will respond. This is very different from the system in England, where students have to study canonical texts with little choice for the teachers as to what they are. In two of the Ontario schools, for example, the teachers spoke of how they used to do *The Fifth Business* by Robson Davies [37] and had now changed because students did not like the novel. In Millennium, they had recently changed and now taught *Indian Horse* by Richard Wagamese [38], about a young, First Nation character, who was abused in the care home system.

It is phenomenal and the kids connect with it and the reluctant reader boys and I have a student who wrote his final reflection on this that he is half Mohawk and that his whole life he would lie and tell people he was Hispanic because he never heard anything good about being Indian. Then he read this book, it was the first time he ever felt proud to say I'm a Mohawk (Marian).

The desire for them to identify with the text is very strong and links Marian with the personal growth model of teaching, [39] which is found in the Ontario curriculum [18].

As a creative representation of life and experience, literature raises important questions about the human condition, now and in the past. As students increase their knowledge of accomplished writers and literary works, and vicariously experience times, events, cultures, and values different from their own, they deepen their understanding of the many dimensions of human thought and human experience [18] (ibid., p. 16).

Indian Horse does this for Marian:

Two years ago that kid wouldn't have had that book . . . Right, so this is the darkest mark on our history and we had kids who hadn't even covered it in history class, even when they took the 20th Century Canadian history class, it was news to them in Grade 12 that this was a thing that had happened in Canada. Astonishing (Marian).

Yet there are problems with the way Ontario chooses to assess through coursework based on what students complete in school. There is no means of telling whether one teacher or a school is in any way tougher or more lenient than their peers. There is nothing on the reliability or validity of the way a teacher assesses so that the grading in one school may differ fundamentally from that of another. It is entirely possible that any given school could attract a reputation as either a school that reliably and validly assesses or as one that is over generous. This might affect perceptions of the school by parents, or perceptions of the students that graduate by university admissions tutors. Even where a system apparently runs on a sense of professional trust, it is likely that there are those that will still ask questions. Effective moderation processes are often seen, understandably, as critical in any system that runs on a purely teacher assessed system.

Other countries or states, as we have noted, introduced some form of moderation into the assessment process. So for example, Queensland, until recently, also operated a system of 100% coursework based on the Radford Report [40] in which they had rigorous moderation [41,42] that was regularly researched [43–45].

Several attempts were trialled in the United States, particularly in the 1990s. Thomas et al. [46], Le Mehieu et al. [5], and Stecher [47] all researched ways of assessing students' work through portfolios rather than timed tests. Each agreed on the benefits; all thought some form of moderation important. When 100% coursework was permitted as a mode of assessment in England, starting in the early sixties, the then Joint Matriculation Board introduced a review panel of experienced teachers who assessed students' work alongside individual teachers marking their own pupils' assignments. It was a way of externally moderating as well as internally moderating, students' work. They found that, 'High correlations between schools and moderators were achieved' [48] (p. 10). Cross moderation was also carried out with the new, mode 3, Certificate of Secondary Education. As the GCSE developed in England it was common for examination boards each year to publish a selection of exemplary student work; teachers in individual departments would agree on marks for each piece of work, and the department heads would meet at a local level as part of a standardisation process. With the standard set, it was then common for English departments in schools to internally moderate coursework each year before submitting marks to the examination boards. This was clearly time-consuming and costly, but it addressed the issues of reliability that can otherwise be aimed at coursework assessment. It is interesting then that the system in Ontario does not require any form of moderation at all.

It is true that in some schools individual teachers meet up and compare grades, as they did in Marian's school yet, 'There isn't a more, that is voluntary and teacher initiative based and there isn't a more formal, every teacher from the Board will come and participate' (Marian). Although her school board, 'Is moving towards a central moderated marking model, where they brought teachers from every school to a marking session, provided us with the anchor papers, the Rubrics' (Marian) it had not actually happened and even this was not a system of extensive moderation, it was just within one school board.

Having said that, the International Baccalaureate is become increasingly popular in Ontario and Canada in general, because it is 'Instilling confidence in [university] admission officers' [49]

(p. 21) and that unlike the qualifications that students usually have there was, ‘The rigour of the programme’ [49] (p. 21). There is coursework for the assessment of the IB, but this is externally, as well as internally, moderated. Nevertheless, ‘The high stakes and high accountability nature of testing’ even the moderating of coursework, ‘That gives prominence to a narrow set of outcomes . . . tend to distort learning and teaching’ [1] (p. 70) clearly evidenced in England that is not present in Ontario. Nor is it there in Saskatchewan.

5. Formative Assessment

The nature of summative assessments can also affect the kind of formative assessment or assessment for learning (AfL) that is practised in schools [50,51].

5.1. England

It is now twenty years since the original *Inside the Black Box* publication [52] brought the notion of assessment for learning to teachers’ collective consciousness. It is fair to say that over the course of the past two, nearly three, decades assessment for learning strategies—questioning techniques, self and peer assessment and the like—have become part of the framework of a huge number of secondary English classrooms in England. One might even argue that the explicit dialogue around exam criteria and assessment foci has some basis in ideas about assessment for learning, although this might be to seek a spurious justification for an approach that perhaps looks more like ‘teaching to the test’. Mary James pointed out that,

Assessment for learning’ is becoming a catch-all phrase, used to refer to a range of practices. In some versions, it has been turned into a series of ritualised procedures. In others, it is taken to be more concerned with monitoring and recording than with using information to help learning [53] (p. 2).

Although it is true that AfL can operate under any system of assessment, be it terminal exams or course-based summative assessment the differences in the way AfL was practised differed starkly. It certainly was the case that in the lessons we observed what might be called assessment for learning techniques featured relatively heavily. Certainly, the whole class dialogue in all the lessons, to a greater or lesser extent, saw teachers targeting questions and inviting students to comment on or develop the responses of their peers. This kind of approach, which is clearly different to the initiation-feedback-response (IFR) model of the whole class talk, allows an opportunity for a teacher to formatively assess. Lisa’s comments about the way she structured whole class feedback at the end of an activity clearly highlighted elements of AfL at work:

I suppose that’s part of the walking around and working out who it is that I want to feedback. I suppose as well it’s using kind of like the positive appraisal, kind of recognising the work that they’ve done that’s good in the lessons. I do think that’s often the best way to do the feedback because if you do that whole hands up thing, you always get the same people that are really reluctant to put their hands up and there are often some students in there who will come up with something really good, but they are not going to share that unless I kind of prompt that they have done it.

Even if it can be legitimately claimed, however, that the practice of making assessment criteria explicit is actually an assessment for a learning strategy, there was no real sense in the lessons that the sharing of the criteria was done with a genuine aim of encouraging students to understand or ‘own’ the assessment frameworks. Although in two of the lessons there were peer assessment activities where students were asked to comment on the writing of a classmate to identify where and when they had ‘hit’ the assessment criteria in their written responses, these activities did not seem to be particularly formative in any way. The feeling was that students were being asked to ‘tick off’ objectives where they saw them and that hitting the objectives was a measure of a successful answer—whether what had

been written was of any real merit, or how it might have been improved, was not really the point of the exercise, it seemed. Marshall and Drummond [28] have written about the 'spirit' and the 'letter' of assessment for learning; it certainly seemed to a large extent in these lessons that whilst AfL techniques were used these were strategies to serve the demands of the external examination rather than part of a learning and teaching approach. Here then, it was the 'letter' of AfL, rather than its spirit as part of a classroom learning culture, that was at work.

The notable exception to this was Paul. In an hour-long lesson, he only mentioned that they were preparing for an examination once. He too subscribes to the view that English is constrained by the examination system, claiming that 'You are constantly following after something that you are not actually sure what you are following after'. He also says 'you are not doing what you should be doing in your job', and that,

I think that endlessly doing something because you have to do it rather than because you want to do it and because the kids are interested in it is quite deadly to the study of English. There are only so many times one can underline an imperative verb and ask someone to write a rhetorical question before it loses all meaning.' (Paul).

So he teaches the way in which he believes English should be taught rather than how he thinks the exam system demands. His lesson on Golding's *Lord of the Flies* [54] was striking in the dialogic nature of the class. It might be said that when speaking and listening happens, it is the first of that pair that attracts the focus, but in this lesson it was, in fact, the quality of Paul's listening that seemed to make the difference; his gaze, facial expressions and body language when pupils spoke conveyed a genuine interest in what was said, which undoubtedly encouraged pupils to say more and to be more forthcoming and explorative in their comments. And the fact that the pupils' comments were listened to by the teacher rather than the teachers' questions, that drove the direction of the talk and gave the pupils a sense that what they were saying was of value, and was critical in moving the discussion on. It would not be too much to suggest that this gave the class a sense of ownership of the lesson content and a responsibility for the lesson trajectory.

It also meant that the questions Paul asked were cumulative. In other words, it was a kind of feedback that genuinely built on the responses that the pupils had given. Alexander [55], talks of the nature of the of dialogic teaching and compares it to AfL saying that Black and Wiliam talk of the importance of dialogue in the classroom in a way 'That [is] remarkably similar to those we advance here for dialogic teaching, to the extent that we might offer the alternative term dialogic assessment' [55] (p. 33).

The turns of the conversation were not predetermined but did not mean that the talk was directionless; the way in which Paul synthesized comments from a range of pupils and used these to formulate subsequent questions made sure this was not a meandering discussion that ran the risk of disappearing up conversational blind alleys. It was clear by the lesson's conclusion that the pupils' thinking had been taken some significant way forward in terms of their understanding of characterization and plot. Also, however, it appeared that their sense of themselves as readers and critics of literature more widely had evolved in a way that means what was happening was more than simply the learning of a text for an exam or piece of coursework.

To give an indication of Paul's techniques we can consider the opening section of the lesson. After a teacher reading of a section of *Lord of the Flies*, the pupils were invited to consider the ways in which the relationship between Jack and Ralph, two central characters, had changed over the course of the text. Paul simply invited the pupils to discuss in pairs 'Anything you have to say about the relationship', a genuinely open invitation for thought and talk, and after a period of ten minutes pupils were invited to contribute to the whole class feedback with the similarly open 'So, what have you to say about the relationship?'

Interestingly, when questioned about the very open-ended way in which he directed the discussions, Paul indicated that this was influenced by the complexities of the text:

... The great thing about *Lord of the Flies* is that it is utterly ambiguous, it's totally horrific and yet at the same time entirely realistic and we don't want to believe these things are possible and yet we can all see the evidence in the world around us that suggests that human beings are capable of terrible things. I think when you approach the text with a particular viewpoint in mind over something as ... trivial is not quite the right word, but something as secondary as who is the most effective leader and why, you are probably doing the text an injustice to say well Ralph is because he is this, this, this and this, surely the point is that there is good and evil in everyone and that I suppose was the theme that I was guiding them towards at the end of the text.

Paul's celebration of the ambiguity of *Lord of the Flies* was evident in the ways he listened and responded to the pupils so that there was a genuine sense that multiple readings were being encouraged and that there were no easy answers.

5.2. Canada

Again the inclusion of assessment for learning in Ontario is very different. Indeed, the term AfL was only mentioned by Maria in Millennium.

When I started teaching 12 years ago, the big PD [professional development] move was the three forms, the assessment for learning, assessment, as learning assessment of learning. Yes, that has been I would say like our Board's guiding framework for the last sort of ten years (Marian).

For Wanda, the origin of the strategies she used come from a very different source yet could still be considered AfL techniques. She used lollipop sticks with pupils' names on them to decide to whom to ask questions and later passed a question on to another student who did not want to speak. She had not heard of these as assessment for learning (AfL) techniques—'think time', 'hands down questioning', and 'phone a friend' [52] Dialogue, and with it questioning, lies at the heart of assessment for learning: 'All such work involves some degree of feedback between those taught and the teacher, and this is entailed in the quality of the interaction which is at the heart of pedagogy' [56] (p. 16). Yet talking about why she used this technique she said of the class,

So their job in the group is to get knowledge based on whatever task I've given to them and so ideally that I'm going to pull your name from this cup, keeps everyone sort-of on their toes, but isn't super risky because everyone has had the opportunity to review the material (Wanda).

Significantly she 'would never do that in a situation where they had not been partnered to talk about something first' (Wanda). The reason she used the strategy, however, was that she had learned it from, 'One of my professors at ... 20 years ago' (Wanda.). She went on to say that he had,

Taught a class called 'Cooperative Learning' and um, it was full of strategies to engage students in those kinds of ways. So to ensure you've got a room full of people who know that they are required to have the information and share it. So that is something that is one of the techniques that I like, I think it's from him (Wanda).

The idea is similar to that of AfL 'hands down' where the teacher can ask anyone a question and so, to quote Wanda, it 'keeps everyone sort-of on their toes'. Yet the idea of 'wait time', where she 'would never do that in a situation where they had not been partnered to talk about something first' (Wanda); passing on a question, and indeed 'hands down questioning', arose not so much from AfL but from the notion that you had to have 'co-operative learning. For Black and Wiliam [57], part of the theoretical underpinning of AfL comes in part from Vygotsky and activity theory, both of which emphasise communities of practice.

It is in Marian's class, however, that the use of AfL is almost more significant. She used formative assessment during a 'culminating' task, a task 'with numbers on them' (Marian). Students had to complete a formal presentation of the work that they had undertaken helping a fifteen-year-old student with literacy. Marian sat at the back of the class and intermittently asked questions of the students. In so doing her feedback was formative in that, through the questions she asked she improved the student's answers [58]. Again they were cumulative. When asked why she had done this in what was meant to be a culminating evaluation she explained that,

It's just, in some cases looking at the Rubric they were given, I could see that they were in the neighbourhood of a point and wanted to give them an opportunity to . . . To hit the target more by posing the questions specifically, because it's the first day of presentations. They are a bit nervous and they would feel bad if they had a point in their mind and they just for some reason didn't quite articulate it in a way that made sense to me. So feeding them the leading question and then my hope is that with the students who are observing because it's the first day of the presentation will take it into account.

Not only was the formative feedback aimed at the particular student doing his presentation, but it was also aimed at the rest of the class too in order that they all might improve. The students were in their final year at school and so this presentation would be one of the contributory assessments in their entrance into university. Such a contribution by a teacher would not only be impossible in England, but it may well be considered, by some, cheating.

6. Conclusions

Ours is only a very small sample of teachers, through our experience of teachers outside these case studies in England is greater than that of Canada. It is not, however, possible to make any generalisations, as our data is small. Yet, the way in which assessment plays a role in how the teachers, in our case studies, teach is significant. In England, the thought of exams, the results, and league tables dominate the way in which teachers approach their lessons. So concerned are they with the summative that even the formative has the examination as the target rather than improving the quality of the students' English, be it spoken or written. It is true that Paul proves an exception to this and in this respect shows that the spirit AfL can and does take place within a system of terminal exams. In resisting pressure, he demonstrates that what goes on in English classrooms need not necessarily bend to the will of policymakers when policy is anathema to what teachers' professional knowledge and experience tell them. But it takes supreme confidence to resist. Yet, if Paul's lesson was not typical, it was at least evidence that even within tight constraints there are ways for teachers to assert a level of control and autonomy. And the constraints are tight. In our case-study sample of teachers in Ontario, by contrast, teachers are allowed an autonomy and this impacts both how they consider what they teach and the impact this has on their classrooms.

This does not mean that England must copy Canada. It is easy to make bold claims based on comparisons. The ranking of Canada sixteen places above England and nineteen above the UK in general screams that the way English is taught in England ought to be jettisoned for one that is more like the Canadian model. Yet Dylan Wiliam's sage point that 'Looking at what other countries are doing has its place, but the solutions will have to be home-grown' [14] (p. 90) ought to be acknowledged. The system of assessment in Ontario may seem blissful but there are drawbacks and it may well be beginning to change. The accountability culture in England might prove difficult but it would be hard now to dismantle completely. It may be, however, that we need to rethink the sharp divide, 'Between objective and subjective judgement . . . the former routinely associated with standardised testing, and the latter, teacher judgement' and acknowledge that such a divide is an 'ill-conceived notion' [1] (p. 68).

Author Contributions: B.M. conceived the premise for the article and the methodology; undertook a substantial amount of the analysis and oversaw the writing, reviewing and editing of this paper. S.G. contributed some of the analysis and writing up.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Klenowski, V.; Wyatt-Smith, C. The impact of high stakes testing: The Australian story. *Assess. Educ. Princ. Policy Pract.* **2012**, *19*, 65–79. [[CrossRef](#)]
2. Mellon, J.C. *National Assessment and the Teaching of Writing: Results of the First National Assessment of Educational Progress in Writing*; National Council of Teachers of English: Urbana, IL, USA, 1975.
3. Freedman, S.W. (Ed.) *Evaluating Writing Linking Large Scale Testing and Classroom Assessment*. In *Centre for the Study of Writing, University of California Berkeley Occasional Paper 27*; Office of Educational Research and Improvement: Washington, DC, USA, 1991.
4. Le Mahieu, P.G.; Gitomer, D.H.; Eresh, J.T. Portfolios in large-scale assessment: Difficult but not impossible. *Educ. Meas. Issues Pract.* **1995**, *14*, 11–28. [[CrossRef](#)]
5. Koretz, D. Large-scale Portfolio Assessments in the US: Evidence pertaining to the quality of measurement. *Assess. Educ. Princ. Policy Pract.* **1998**, *5*, 309–334. [[CrossRef](#)]
6. Wiliam, D. *Assessment for Learning: Why no profile in US policy*. In *Assessment and Learning*; Gardener, J., Ed.; Sage: London, UK, 2009.
7. Dohn, N.B. Knowledge and Skills for PISA—Assessing the Assessment. *J. Philos.* **2007**, *41*, 1–16. [[CrossRef](#)]
8. Bulle, N. Comparing OECD educational models through the prism of PISA. *Comp. Educ.* **2011**, *47*, 503–521. [[CrossRef](#)]
9. Pereyra, M.A.; Kotthoff, H.G.; Cowen, R. (Eds.) *PISA under Examination*; Sense Publishers: Rotterdam, The Netherlands, 2011.
10. Murphy, D. Issues with PISA’s Use of its Data in the Context of International Education Policy Convergence. *Policy Futur. Educ.* **2014**, *12*, 893–916. [[CrossRef](#)]
11. Jerrim, J. The reliability of trends over time in international education test scores: Is the performance of England’s secondary school pupils really in relative decline? *J. Soc. Policy* **2012**, *42*, 259–279. [[CrossRef](#)]
12. Breakspear, S. The policy impact of PISA: The exploration of the nominative effects of international benchmarking in the school system performance. *OECD Education Working Paper*. **2012**. Available online: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.463.3776> (accessed on 12 November 2018). [[CrossRef](#)]
13. Andrews, P.; Atkinson, L.; Ball, S.J.; Barber, M.; Beckett, L.; Berardi, J.; Zhao, Y.; Berliner, D.; Bloom, E.E.; Boudet, D.; et al. OECD and PISA Tests Are Damaging Education Worldwide—Academics. Available online: <https://www.theguardian.com/education/2014/may/06/oecd-pisa-tests-damaging-education-academic> (accessed on 6 May 2014).
14. Wiliam, D. *Creating the Schools Our Children Need: Why What We’re Doing Right Now Won’t Help Much, and What We Can Do Instead*; Learning Sciences International: West Palm Beach, FL, USA, 2018.
15. Oates, T. Could Do Better: Using international comparisons to refine the national curriculum in England. *Curric. J.* **2011**, *22*, 121–150. [[CrossRef](#)]
16. Thomas, S.; Gana, Y.; Muñoz-Chereau, B. England: The intersection of international development and educational policy development. In *The Intersection of International Achievement Testing and Educational Policy: Global Perspectives on Large-Scale Reform*; Volante, L., Ed.; Routledge: London, UK, 2016.
17. Young, J.; Boyd, K. More than servants of the State? The governance of initial teacher-preparation in Canada in an era of school reform. *Alta. J. Educ. Res.* **2010**, *56*, 1–18.
18. Ministry of Education. *The Ontario Curriculum Grades 9 and 10: English*; Queen’s Printer for Ontario: Toronto, Canada, 2007.
19. Gillham, B. *Case Study Research Methods*; Continuum: London, UK, 2000.
20. Gillham, B. *The Research Interview*; Continuum: London, UK, 2000.
21. Gillham, B. *Research Interviewing*; Open University Press: Maidenhead, UK, 2005.
22. Stake, R.E. *The Art of Case Study Research*; Sage Publications: Thousand Oaks, CA, USA, 1995.

23. Bassey, M. *Case Study Research in Educational Settings*; Open University Press: Buckingham, UK, 1999.
24. Elger, T. *Encyclopedia of Case Study Research*; Mills, A.J., Durepos, G., Wiebe, E., Eds.; SAGE Publications: Los Angeles, CA, USA, 2010.
25. Swanborn, P.G. *Case Study Research*; SAGE Publications: Los Angeles, CA, USA, 2010.
26. Marshall, B.; Gibbons, S. A Methodological Conundrum: Comparing Schools in Scotland and England. *Changing English: Studies in Culture and Education* **2015**, *22*, 199–208. [CrossRef]
27. Marshall, B.; Pahl, K. Who Owns Educational Research? Disciplinary Conundrums and considerations. *Qual. Res. J.* **2015**, *15*, 472–478. [CrossRef]
28. Marshall, B.; Drummond, M.-J. How Teachers Engage with Assessment for Learning: lessons from the classroom. *Res. Paper Educ.* **2006**, *21*, 133–154. [CrossRef]
29. Marshall, B.; Gibbons, S.; Hayward, L.; Spencer, E. *Policy and Practice in the Secondary English Classroom: A Case Study Approach to the Teaching of Secondary English in Canada, England and Scotland*; Bloomsbury Publishing: London, UK, 2018.
30. Marshall, B. *English Teachers—The Unofficial Guide: Researching the philosophies of English teachers*; Routledge Falmer: London, UK, 2000.
31. Eisner, E. *Reimagining Schools: The Selected Works of Elliot W. Eisner*; Routledge: London, UK, 2005.
32. Dewey, J. *Art as Experience*; Perigree: New York, NY, USA, 1932.
33. Department for Education. National Curriculum in England: English Programmes of Study. Available online: <https://www.gov.uk/government/publications/national-curriculum-in-england-english-programmes-of-study/national-curriculum-in-england-english-programm> (accessed on 17 March 2018).
34. Department for Education. The Framework for the National Curriculum: A Report by the Expert Panel for the National Curriculum Review. Available online: http://www.gov.uk/government/uploads/attachment_data/file/175439/NCRExpertPanelReport.pdf (accessed on 12 June 2016).
35. Steinbeck, J. *Of Mice and Men*; Penguin: London, UK, 2000.
36. Office for Standards in Education. *Moving English Forward: Action to Raise Standards in English*; Ofsted: London, UK, 2012.
37. Davies, R. *Fifth Business*; Penguin: London, UK, 2015.
38. Wagamese, R. *Indian Horse*; Douglas McIntyre: Vancouver, BC, Canada, 2012.
39. Gibbons, S. *English and Its Teachers: A History of Policy, Pedagogy and Practice*; Routledge: Oxford, UK, 2017.
40. The Radford Report Queensland. *Committee Appointed to Review the System of Public Examinations for Queensland Secondary School Students and to Make Recommendations for the Assessment of Student's Achievements*; Department of Education: Brisbane, Australia, 1970.
41. Queensland Studies Authority. *P–12 Assessment Policy*; QSA: Brisbane, Australia, 2009.
42. Queensland Studies Authority. Available online: <http://qsa/qld.edu.au/assessment/586.html> (accessed on 7 January 2010).
43. Maxwell, G. Progressive Assessment for Learning and Certification: Some Lessons from school-based assessment in Queensland. In Proceedings of the Third Conference of the Association of Commonwealth Examination and Assessment Boards Redefining the Roles of Educational Assessment, Nadi, Fiji, March 2004.
44. Klenowski, V.; Adie, L. Moderation as judgement practice: Reconciling system level accountability and local level practice. *Curric. Perspect.* **2009**, *29*, 10–28.
45. Klenowski, V.; Wyatt-Smith, C. Standards-driven reform Years 1–10: Moderation an optional extra? In Proceedings of the Australian Association for Research in Education Conference, Brisbane, Australia, 2–6 December 2008.
46. Thomas, W.H.; Storms, B.A.; Sheingold, K.; Heller, J.J.; Paulukonis, S.T.; Nunez, A.M.; Wing, J.Y. *California Learning Assessment System: Portfolio Assessment Research and Development Project: Final Report*; Educational Testing Service, Center for Performance Assessment: Princeton, NJ, USA, 1995.
47. Stecher, B. The Local Benefits and Burdens of Large-scale Portfolio Assessment. *Assess. Educ. Princ. Policy Pract.* **1998**, *3*, 335–351. [CrossRef]
48. Wilson, F.; Hewitt, E.A.; Gordon, D.I. *English Language: An Experiment in School Assessing: First Interim Report*; Joint Matriculation Board: Manchester, UK, 1965.
49. Fitzgerald, S. Perceptions of the International Baccalaureate (IB) in Ontario Universities. *Can. J. Educ.* **2015**, *38*, 1–34.

50. Black, P.J.; Harrison, C.; Hodgson, J.; Marshall, B.; Serret, N. Validity in teachers' summative assessments. *Assess. Educ. Princ. Policy Pract.* **2010**, *17*, 215–232. [[CrossRef](#)]
51. Black, P.; Harrison, C.; Hodgen, J.; Marshall, B.; Serret, N. Can teachers' summative assessments produce dependable results and also enhance classroom learning? *Assess. Educ. Princ. Policy Pract.* **2011**, *18*, 451–469. [[CrossRef](#)]
52. Black, P.; Wiliam, D. *Inside the Black Box*; King's College London: London, UK, 1998.
53. James, M. Assessment of Learning: Assessment for Learning and Personalised Learning. In Proceedings of the Goldman Sachs UK/US Conference on Urban Education, London, UK, December 2004.
54. Golding, W. *Lord of the Flies*; Faber and Faber Limited: London, UK, 2004.
55. Alexander, R.J. *Towards Dialogic Teaching Rethinking Classroom Talk*, 4th ed.; Dialogos UK Ltd.: North Yorkshire, UK, 2010.
56. Black, P.J.; Harrison, C.; Lee, C.; Marshall, B.; Wiliam, D. *Assessment for Learning: Putting It into Practice*; Open University Press: Buckingham, UK, 2003.
57. Black, P.; Wiliam, D. Developing the Theory of Formative Assessment. *Educ. Assess. Eval. Account.* **2009**. [[CrossRef](#)]
58. Marshall, B. *Testing English: Summative and Formative Assessment in English*; Continuum: London, UK, 2011.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).