The Enactment of Applied English: Does Caring Lead to Teaching to the Test?

Stefan Merchant
Queen's University

Jessica Rich
Queen's University

Don Klinger
University of Waikato

Rebecca Luce-Kapler
Queen's University

Abstract

Many school systems require students to pass one or more standardized examinations to graduate. This case study of three lower academic stream English classes in a single Ontario secondary school explored how the presence of a high-stakes standardized examination affected the enactment of the curriculum. Qualitative data were collected through a combination of teacher interviews, classroom observations, and teaching/learning artifacts. The findings of the inductive analysis suggest that the four teachers in the study cared deeply about the emotional well-being and scholastic success of their students; however, this caring sometimes manifested itself in less than ideal teaching practices, such as the use of external rewards, removal of student responsibility for learning, and teaching to the test. Seven strands are identified, each describing how caring is manifested within this teaching/learning context. Three vignettes describe actual classroom events which epitomise how caring was demonstrated by these teachers towards their students. While positioning teaching to the test as an act of caring is unusual, the evidence from this study suggests teachers do it because they believe it will benefit students' well-being and future outcomes. These findings suggest that for any high-stakes test, caring teachers may teach to the test and narrow the curriculum.

Keywords: English curriculum, standardized testing, caring, washback, motivation, learning skills

Résumé

Nombre de systèmes scolaires exigent que les élèves passent un ou plusieurs examens standardisés pour obtenir leur diplôme. Cette étude de cas, portant sur trois classes d'anglais dans une école secondaire d'Ontario, a exploré l'impact de la présence d'un examen standardisé à enjeux élevés sur la mise en œuvre du programme d'études. Les données qualitatives ont été recueillies au moyen d'une combinaison d'entrevues avec des enseignants, d'observations en classe et d'outils d'enseignement/apprentissage. Les résultats de l'analyse inductive indiquent que les quatre enseignants de l'étude se préoccupaient du bien-être émotionnel et de la réussite scolaire de leurs élèves ; cependant, cette attention s'est parfois manifestée dans des pratiques d'enseignement moins qu'idéales, comme l'utilisation de

récompenses externes, la suppression de la responsabilité des étudiants en matière d'apprentissage et l'enseignement en fonction de l'épreuve. Sept thèmes sont identifiés, chacun décrivant comment se manifeste le dévouement de l'enseignant dans ce contexte d'enseignement/apprentissage. Trois scènes décrivant des événements réels en classe illustrent la façon dont le dévouement a été démontré par ces enseignants envers leurs élèves. Bien qu'il soit inhabituel de considérer l'enseignement en fonction de l'épreuve comme un acte de dévouement, les résultats de cette étude portent à croire que les enseignants le font parce qu'ils pensent que cela sera bénéfique pour le bien-être et les résultats scolaires des élèves. Nos résultats indiquent que pour tout test à enjeux élevés, les enseignants dévoués sont susceptibles d'enseigner en fonction de l'épreuve et de réduire leur programme d'études.

Mots-clés : programme d'anglais, tests standardisés, enseignant dévoué, retour en arrière, motivation, compétences d'apprentissage, enseignement en fonction de l'épreuve

Introduction

Many school systems require students to pass a high-stakes examination for graduation (International Reading Association, 2014). For example, thirteen of the 50 United States require students to pass at least one examination for graduation (Gewertz, 2019) and both British Columbia and Ontario require students to pass a literacy examination to graduate. The Ontario examination is called the Ontario Secondary School Literacy Test (OSSLT) and is a minimum proficiency test written by Grade 10 students. Its purpose is "to determine whether a student has the literacy (reading and writing) skills required to meet the standard for understanding reading selections and communicating in a variety of writing forms expected by The Ontario Curriculum across all subjects up to the end of Grade 9" (EQAO, 2006, p. 4). The pass rate for the OSSLT has been stable over the last five years at about 80% (EQAO, 2017). The 20% failure rate means that approximately 30,000 students have their high school graduation threatened. These students must then retake (and pass) the OSSLT in a subsequent year or complete a supplementary literacy course to graduate.

Most Grade 10 students in Ontario are streamed into either Applied or Academic English. While the overall curriculum expectations for these two courses are similar, students who are perceived as less capable English learners are generally streamed into Applied English. Boys and low socio-economic status (SES) students are also overrepresented in Applied English classes (Card & Payne, 2015; Hamlin & Cameron, 2015). This distinction in learning streams is important because enrolment in Applied English is not only the most significant predictor of passing the OSSLT, but also the one affecting the largest number of students (EQAO, 2017). While over 90% of Academic English students typically pass the test, only 44% of Applied English students are successful (EQAO, 2017).

Compared to some jursidictions, the use of standardized testing in Ontario is limited. In addition to the OSSLT, students write standardized tests in Grades 3, 6, and 9. These examinations involve low stakes for teachers. The results of these tests are not used for grade promotion, tenure decisions, performance bonuses, or formal evaluations of teaching. This is not to say that Ontario school administrators (and parents) do not care about standardized test results, but teachers are not formally rewarded or punished for their class's achievement on the tests. For teachers of Applied English classes, our own

(admittedly informal and limited) experiences in schools indicates there is a belief that teaching Applied English is a noble sacrifice the teacher makes. These are difficult classes, and a teacher's reputation with their administration is determined not by OSSLT pass rates, but by their willingness to teach the course and manage student behaviour.

All assessments have consequences—some of which are intended, and others unintended and detrimental (Stobart, 2003). The minimum competency nature of the OSSLT and the wide discrepancy in pass rates between Academic and Applied English streams means the examination is of little consequence for most Academic English students, but is of significant importance for those in Applied English. Not graduating from high school is an intended consequence of a student failing the OSSLT, but other unintended consequences exist. Prior research suggests that the experience of failing the OSSLT harms students' self-esteem, their relationships with school classes, and their perceptions of their abilities in English (Kearns, 2011). Poor self-esteem is associated with students dropping out of high school, and Canadians who do not graduate from high school find it difficult to obtain employment (Uppal, 2017). Furthermore, high school dropouts are much more likely to rely on government transfers for a portion of their income (Uppal, 2017). The personal and social consequences of not completing high school are enormous. Luce-Kapler and Klinger (2005) also identified perceived consequences with respect to future educational and career choices, as students use the OSSLT as a gauge of their abilities to pursue specific academic or career pathways.

Given the known disadvantages experienced by students failing the OSSLT, we felt it important to ask and answer the question: What are teachers of Applied English doing in schools to help improve OSSLT outcomes for these students? Using a case study of three different Applied English classes in one high school, our research offers insight into how teachers of Applied English enact the curriculum and support student success on the OSSLT. The qualitative data presented here focus on teachers and consists of interviews, classroom observations, and documentary evidence, including student handouts and teacher-created assessments. This rich data set enabled us to begin understanding why teachers implement the Applied English curriculum in the way they do. We chose a case study methodology because case studies are good for answering "how" and "why" questions, and are effective means of studying complex phenomena in context (Merriam, 1998; Yin, 2018).

Caring in the Classroom

As former teachers ourselves, we find that we naturally adopt an appreciative stance toward teachers. Every profession, including teaching, has members who fail to uphold standards of ethical practice, but our view is that most teachers strive to be ethical and professional, and care deeply about their students. However, to say that a teacher is caring is to provide little information about what that teacher does in practice. Thus, we ask, "What does caring look like in the Applied English classroom?" In answering this question, our starting point is Noddings (1984). Her concept of caring arises from Mayeroff's (1971) view that "to care for another person, in the most significant sense, is to help him grow and actualize himself" (p. 1). Noddings's description of caring is slightly different. In her view, caring for someone is having "regard for his views and interests" (Noddings, 1984, p. 9). Both of these descriptions would likely be agreeable to most teachers, but they lack specificity. They fail to describe what teachers do to demonstrate caring in the classroom.

Noddings (1984) addresses this question reluctantly. She writes, "In spite of the difficulties involved, we shall have to discuss behavioral indicators of caring in some depth" (p. 12) and then goes on to discuss "engrossment." Engrossment is described as "apprehending the other's reality, feeling what he feels as nearly as possible" (p. 16). While engrossment may be an essential element of caring, it is not a behaviour or action—it is a stance. In fact, as White (2003) has shown, the literature on caring has surface-level and often contradictory accounts of the teacher behaviours that are markers of caring. For example, Dodds-Urban (1999) characterizes caring as persevering with students, even when they are unreceptive, while Noddings (1984) demands that caring relationships be reciprocal. Indicators of caring in the classroom have included the way a teacher looks at a student (Langer, 2000), pronouncing students' names correctly (Stewart, 2016), and listening to students (Nelson, 2000). These behaviours may be good indicators of caring, but a complete list of indicators would be impossible to compile. How caring is enacted in classrooms is complex and dependent on context (Lee & Schallert, 2008). Whether a behaviour is an act of caring is not always obvious from the behaviour itself. For example, a teacher hugging a student may be caring, or it may be predatory.

As evidenced by the indicators mentioned above, the caring literature consistently conceptualizes caring as being between a teacher and an individual student (e. g., Baker,

2006; Barrow; 2015; Noddings, 1984). Good relationships between the teacher and individual students are important, as well as associated with a host of positive educational and affective outcomes (Yates & Hattie, 2013). However, much of a high school teacher's day is spent interacting not with individuals, but with an entire class, and the caring literature does not tell us, or even address, what a teacher caring for a whole class looks like. What lesson plans or learning activities does a caring teacher use? What does caring assessment look like?

Part of the reason for this omission from the caring literature may be its focus on the affective domain. Noddings (1988) believes that for caring teachers, affective goals outweigh academic goals. As she puts it, "Teachers are, of course, concerned with their students' academic achievement, but more importantly, they are interested in the development of fully moral persons" (p. 222). Noddings rails against "our endless focus on narrow achievement goals" (p. 226), but does not consider that for a high school teacher such a focus may be a deep expression of caring. In a society where entry into post-secondary institutions is governed by academic achievement, and where graduation may hinge upon success on a standardized test, a teacher who is "teaching to the test" may be doing so out of genuine concern for students. Thus, we posit that in a high school classroom, caring may be manifested as teaching to the test.

Washback

Volante (2004) states that "the job of any teacher is first and foremost to promote learning in their students" (p. 1), but what if this is not how the teacher sees the job? Barnes's (2017) case study of four English teachers found that in the context of a Test of English as a Foreign Language (TOEFL) preparation course, the teachers saw the purpose of their teaching as improving test scores—not improving English language skills. Given that students' future opportunities can hinge on a standardized test, such as the TOEFL, teaching to the test may be done because the teacher views it as being in the best interests of the student.

Multiple studies conducted over decades have found that high-stakes tests affect both what (i.e., content) and how (i.e., methods) teachers teach (e.g., Alderson & Hamp-Lyons, 1996; Barnes, 2017; Mendelsohn, 1989; Spratt, 2005; Wong, 1969). These results are found not only in North America, but globally (e.g., Damankesh & Babaii,

2015; Kilickaya, 2016). One term used to describe the impacts that standardized testing can have on teaching and learning is washback (Tsagari & Cheng, 2016). Researchers have described a variety of negative consequences that would be considered washback. Popham (2001) notes that training students to respond to specific test items "eviscerates the validity of score-based inferences" (p. 17), while Berliner (2011) says that high-stakes testing can lead to a narrowing of the curriculum such that the test becomes the curriculum. Others have argued that teaching to the test results in too many instructional hours being devoted to test preparation instead of meeting curricular expectations (Musoleno & White, 2010; Shepard & Dougherty, 1991), and that large-scale tests emphasize low-level knowledge and recall type questions, resulting in a "dumbing down" of what is being taught in classrooms (Kohn, 2000).

While various models of washback exist (e.g., Bailey, 1996; Shih, 2007), we have chosen to use Green's (2007) model of washback. In Green's model, washback is influenced both by perceptions of the importance of the test and its difficulty. In this model, washback effects occur when the test is perceived as important and the strength of these effects is moderated by the perceived challenge of the test. If the test is perceived as either very easy or impossible, the washback effects are reduced, whereas they are increased if the test is considered challenging but passable. This model predicts that washback effects are likely to be very high in Ontario Applied English classrooms. Not only is passing the OSSLT required for graduation, but the test is challenging (thoughnot impossible for these students), with success rates typically hovering around 44% (EQAO, 2017).

Study Description/Methods

The results presented here are from our initial case study conducted in a school located in a small Ontario city (population < 50,000). The student body is almost exclusively white, with a mix of low and middle SES students. There were no students in any of the classes who did not have English as their first language. After securing ethics approval from our institutional Research Ethics Board and the District School Board, we obtained permission from the principal of the school to work with four teachers assigned to three different classes (two Grade 10 Applied English and one Grade 9 Applied/Academic English split).

While we would have liked to include more classes, only one school responded positively to our request to conduct research.

Consistent with best practices in case study research, we used multiple methods to collect qualitative data for each case (Stake, 2006; Yazan, 2015). Data were collected using direct observation, documentation, and interviews. The multiple methods of data collection not only provided a rich data set, but also allowed for triangulation, enhancing confidence in the findings (Leech & Onwuegbuzie, 2007; Yin, 2018). Four classroom observations were conducted by members of the research team for each class over a 5-week period. Classroom observations did not focus on evaluating teacher competence, but rather on how teachers enacted the Applied English curriculum (i.e., instructional and assessment practices). During the observation periods, written notes were taken describing classroom schedules, activities, and teacher behaviours. Either during or after observation periods we collected documentation related to the lesson. This documentation included student handouts, worksheets, assignment/task rubrics, and other classroom-based assessments (e.g., tests, quizzes).

Because it is critical to understand teachers' thinking, which guides their lesson planning and delivery, we conducted three semi-structured interviews with each teacher. The first interview took place before the first observation and explored teachers' conceptions of literacy in the context of the Applied English course. During this interview, teachers explained what they saw as the goal(s) of the Applied English curriculum. The second interview occurred after the second observation and asked teachers to describe what curriculum elements they saw as beneficial or harmful to Applied English students. The Ontario English curriculum for Grades 9 and 10 (Ontario Ministry of Education, 2007) was used as an elicitation tool during this interview. The final interview occurred after the final observation and focused on the OSSLT and its perceived impact on students. This interview was also an opportunity for us to ask clarifying questions and for the teacher to contribute additional comments. We adopted what Yin (2018) calls a closed case study design (meaning that all interview and observation protocols were set ahead of data collection and remained constant). While there can be benefits to modifying case study designs as new information emerges, our use of a closed design ensured consistent data were collected across cases, facilitating comparative analyses.

Qualitative analysis of the data began by transcribing the interviews verbatim. Two transcripts were independently and inductively analyzed to establish a set of codes to

use for the other transcripts. Once the initial round of inductive coding had been completed, the two coding authors met to compare codes and resolve disagreements. Initial agreement was 77%. The majority of the remaining disagreements were easily resolved through discussion. The most common source of disagreement arose when it was possible to apply more than one code to a datum. For example, it was possible to interpret test preparation drills as both "teaching to the test" and "repetition." After discussion, the level of agreement rose to 95%. Where agreement on a code could not be quickly reached, that datum was excluded from the data set. This process resulted in an initial set of codes to be used. The same two authors then independently coded two transcripts from another teacher of a different Applied English class. After coding these transcripts, the coding scheme was further refined. Once the final coding scheme was in place, both authors independently coded the remaining transcripts. Coding disagreements at this stage were minimal and resolved through ongoing discussion.

The next phase of coding focused on the collection of artifacts produced by the participating teachers. These artifacts included in-class learning activities, worksheets, and assessments used during the observed lessons. Document analysis was conducted with the same codes used for the interview data. This was done by one of the coding authors and reviewed by the second as a quality check. Finally, the classroom observation notes were coded using the same approach as was used with the artifacts.

Findings

Our findings are first presented as a series of three vignettes representing actual classroom events that epitomize the nature of our observations and highlight critical themes in the data. Vignettes provide the reader with a sense of each classroom and each teacher's approach to enacting the Applied English curriculum (Stake, 1995, 2006). The vignettes are followed by a description of seven strands (themes) we saw emerging from our data. Each strand is described and accompanied by supporting evidence from our data.

Vignette 1: Ms. D, Grade 9 Applied/Academic English Split

The bell rings to begin the second period. Ten of the 17 students are in class, and of these 10, there is an even divide between those napping heads-down, hoods-up at their desks,

and those yelling at peers across the room. Ms. D enters the room and picks up from where the Occasional Teacher left off the previous day—with classroom management. "How dare you destroy the property of my room! I know who it was and I hope that you never do it again because I am *embarrassed*. Everybody up and out of your seats and stand at the back of the room. You are getting a new seating plan." After strategically reseating the class, Ms. D returns to each student the newspaper article they had started—but not finished—on Wednesday. Students had watched the 1992 film *School Ties* and were to write a one-paragraph news report about a key event in the story. Ms. D gives students 30 minutes to finish self- and peer-editing their brief news reports. However, after being bombarded with questions about the plot, Ms. D realizes many students have not written their paragraph and do not remember what happened in the film.

Students rewatch key segments of the film while Ms. D circulates throughout the room, conferencing with individual students about the task, the status of incomplete assignments, and their pending progress reports. After speaking with a few students and scanning their paragraphs, Ms. D realizes that students are not addressing the five Ws (who, what, where, when, and why), and are not including the required direct quote from one of the characters. Ms. D marches to the blackboard and interrupts the work period to review the structure of one-paragraph news reports, reminding the class, "You start writing the paragraph with the most important details and then continue with less important information." As soon as she finishes, students' voices rise again with several questions—many related to spelling: "How do you spell decision?" "How do you spell incident?" After sounding out the words, Ms. D focuses on one student's paragraph, thinking aloud as she reads. "Can I show you some things I would like you to add?" she asks. "You need to add a quote from one of the boys who made a decision in the tribunal." Students are already walking out the door as the bell sounds, and Ms. D warns, "I need this newspaper report from you now or first thing Monday morning. If you don't have at least two of these things handed in [points to the board listing assignments and due dates] it won't be good."

Vignette 2: Mr. C, Grade 10 Applied English

Of the 10 students in this all-male class, five have their binders open and a pen in their hand when the bell rings on this Friday morning. Students who are prepared for class

are rewarded with a ticket to enter the weekly draw for a bottle of pop. Late students are subjected to kibitzing from their peers and occasionally from Mr. C himself. As always, the class agenda is on the blackboard, and the activities follow a pattern. Every period begins with independent work focused on common grammar and punctuation mistakes. On Fridays, this independent work "counts" as a quiz and students complete the work-sheet by "fixing the errors." When orally taking up the work, students receive individually wrapped candies for answering correctly. Wrong answers typically generate more "razzing" from peers.

Next, Mr. C leads the class in a vocabulary building activity, using pictures to elicit vocabulary words from students. After each new word, students are prompted to match each word to its definition and to draw a quick sketch as a visual aid to jog their memory of the meaning. As students finish their sketches, Mr. C reminds the class they have been reading short stories about heroes and asks, "Who can give me two parts of the heroic pattern?" As students provide answers, Mr. C records the answers to the fill-in-the-blank worksheet projected on the SmartBoard. Noticing that he is losing students' attention, Mr. C addresses one student: "Tom—how's the focus?" This prompts Mr. C. to transition to an activity where students complete a Buzzfeed-style quiz: "Which Greek god are you most like?" Once finished, students are instructed to write a paragraph about the Greek god or goddess they would pray to and why. Since the class is slow to transition, Mr. C reminds everyone to follow the handout/guide on paragraph writing and provides prompts for individual students to *start* writing. As the bell rings, Mr. C warns the class: "Your open book test could be difficult next week if there is nothing written in your notebooks."

Vignette 3: Ms. B and Mr. A, Grade 10 Applied English

The second period bell is about to ring. Mr. A is scanning the room looking for specific students who appear to be absent—again. Mr. A makes a habit of asking the class if they have seen or know why students might be absent each day. Class *always* starts on time and with a PowerPoint slideshow. Today, the focus is on inferencing. One by one, Mr. A presents individual paragraphs to the class, asking students to make an inference about the meaning and to raise their hand to share their answer. One student reluctantly responds when called upon. After sharing an incorrect answer, another student is told to

"give it a shot." After the whole class discusses a few examples, Mr. A divides the class into two groups of six and announces there will be a "contest for cake!" Each group is given a sheet containing 16 short paragraphs. They are to infer the meaning of each paragraph and the students are told that Mr. A will bake a cake for the group with the most correct answers. Students are reluctant to work within their assigned group of six; instead, most complete the worksheet with a friend. To promote collaboration, Mr. A instructs the pairs to come together as a group of six to decide on a final set of answers. Groups form, but only after much cajoling. After 10 minutes, the groups dissolve, students go back to their individual seats (organized in single rows), and Mr. A reveals the answers.

For the next activity, students retrieve a copy of the novel *Tex*, by S. E. Hinton, and their duotangs, which are stored safely on the back bookshelf. Mr. A reviews what happened last class and encourages students to follow along as he reads from the book. Periodically, Mr. A pauses to ask comprehension questions and urge someone to take over reading aloud. No one ever volunteers. Sometimes, individual students are told to take over; they never want to, but they comply. At the end of the reading, students know to complete the comprehension worksheet silently in their duotangs and to put their novel and duotang back on the shelf on their way out the door when the bell rings.

Seven Strands

In reading these vignettes, it may be tempting to judge these teachers for motivating students with candies and cake, or for delivering teacher-driven lessons where students hold little responsibility for remembering their materials or directing their own learning. However, these vignettes are not intended to make evaluative judgements about the teachers' competence. Rather, they provide a window into these classrooms and the strategies used to support their students' learning and readiness for the OSSLT. While not always explicit in the classroom observations, our interview data clearly showed a key focus for these teachers was the OSSLT. In recognizing the potential challenges for these students, each of these teachers employed strategies to enhance students' content knowledge, reading skills, and writing strategies in alignment with the OSSLT. They were teaching to the test. However, this is an oversimplification of what was happening in these classrooms, as the teachers were also working to motivate and create personal connections, demonstrating a caring stance for their students. To illustrate this in more detail, we describe seven

strands, which identify manifestations of caring in our data, and give supporting evidence for each.

Caring is knowing the challenges facing their students and trying to save them from "the embarrassment factor" of failing the OSSLT. These teachers were acutely aware of the challenges facing their students, as well as the negative academic and personal consequences of failing the OSSLT. Mr. C, in reflecting on common traits of students in Applied-level courses, described a sense of apathy toward learning and low intrinsic motivation—possibly attributed to parents who "might have not had a good high school experience" and "might not stress achievement at school, or homework." Mr. A reinforced "the fact that many kids who are at the Applied level come to that class with some sort of Individual Education Plan because they have a learning disability… including communication disorders, difficulties with memory," not to mention "a lot of baggage, perhaps some home life and socioeconomic concerns." Ms. B described these students as apprehensive and uncomfortable with taking risks in the classroom "because they are afraid of being wrong, and don't want to embarrass themselves." However, this fear of embarrassment also motivates students to want to pass the OSSLT. According to Mr. C, "They know they need to pass the OSSLT, they would like to do it the first time… So, that is a motivation for many of them. Just so you don't have to tell your friends, 'Yeah, I failed.'"

Note that the focus is not on providing positive emotions, but avoiding negative ones. This is markedly different to Noddings's (1984) belief that caring teachers seek to instill joy in their students. In a description that foreshadows Csikszentmihalyi's (1997) concept of flow, Noddings writes that when someone is totally engaged with a problem, as they struggle and move forward with the problem, they "feel a growing pleasure, a real excitement" (p. 140). This was not the case for the students in the Applied English classrooms we observed. Instead, those students were stuck in a course they did not want to take, reluctantly preparing to pass a mandatory examination. Joy was never part of the classroom experience (at least not joy related to English language and literature), nor did it appear that the teachers in our case study felt that joy was attainable for these students. Thus, caring for students' emotions in these classrooms took the form of minimizing negative emotions, rather than maximizing positive ones.

Caring is controlling behaviour so that learning can happen. All four teachers strongly believed that managing student behaviour was critical to improving learning. Mr. A explained that "when applied classes don't function well, it's often because the teacher has a difficult time managing them. If you can manage them well, then you can get them to learn. But if it's a zoo to begin with, forget it." Similarly, Ms. B commented that "you have to have a highly controlled environment." Even in the Grade 9 Academic/Applied split class, Ms. D prioritized being firm but flexible with her students. "I feel like I just need to run a really tight ship when it comes to behaviour…you have to be tough but kind at the same time."

The most striking example of controlling behaviour from this study was the mid-semester reshuffling of students in the classes. According to Mr. A and Ms. B, the school-level decision to split Ms. B's class and reassign a group of 10 boys to Mr. C "was to accomplish behaviour management, but also to improve the literacy scores." The end result of this reshuffling was that all of the "problem" boys were placed into a single section. This section was given to a teacher (Mr. C) who felt confident he could manage the group. Mr. C described himself as having a genuine affection for the boys, and he also felt that he understood them, and that this understanding helped him manage their motivation and behaviour.

Caring is making Applied English practical and relevant. Each teacher was adamant that Applied English needs to be practical—above all else. As Ms. D explained, "Practical, practical, practical in regards to making sure the ideas and application make sense to students in that they can say, 'OK, I can use this in the future. I will need this.'" According to Ms. D, what is "practical" "are things like grammar, media literacy, how to read a newspaper, how to write a newspaper report," and "being able to make inferences when reading…and making real world applications." Mr. C echoed this point in saying that because these students will go into apprenticeships and the trades, "we don't need to drill them on writing reports or essays." What they need to be able to do is "fill out a paragraph or two for their boss, or a work order form."

When asked to define literacy for students in Applied English, these teachers described the importance of students being "functionally literate," meaning they can "participate actively in their day-to-day lives…and in the world around them as a whole" (Ms. D). In other words, being able to "read and write at an efficient level to be able to survive

in the job force and earn a living" (Ms. B). Key to this aspect of caring was the focus on foundational skills such as vocabulary, comprehension, interpretation, and the ability to communicate one's ideas.

Caring is compensating for students' poor learning skills and work habits. Recognizing that students in Applied English tend to have weak learning skills and work habits, all four teachers compensated for their students' poor organizational skills, lack of initiative, and poor self-regulation. As Mr. A and Ms. B explained, "We don't even let them bring in their own binder. Like it's *that* controlled. We give them a duotang with everything they need in it…it's all right here so I can just say, 'OK, turn to this page X in your duotang' and it stays in the room." Ms. B went on to explain that students in Applied classes "get their work done [in class] but never want to go back over and check their work…no editing, no caps on names, no periods—we see a lot of the same repetitious mistakes." This is why Ms. D frequently provided scaffolded opportunities for students to engage and self- and peer-editing of their writing in class and frequently chased students for their missing work. All four teachers refrained from assigning homework, knowing it would not get done. Experience had demonstrated to these teachers that students in Applied classes would not independently adopt self-regulatory behaviours to support their learning. Hence, these teachers created approaches and strategies to work around these missing skills and reduce the need for them.

Caring is teaching these students how to pass the OSSLT. All three classrooms had an intense focus on OSSLT test preparation and practice. Test preparation included reading and writing skills, vocabulary development, reading comprehension, inferencing, and expository paragraph writing. Writing instruction emphasized grammar, spelling, punctuation, and use of quotations, as these were the types of errors these teachers believed most often caused Applied English students to lose marks on the OSSLT. All four teachers described the need to prepare students "for what the kids need to know on the literacy test. Every year, they have to write a paragraph, a news report, and an opinion piece." Mr. A equated test preparation to coaching, "because you are doing lots and lots and lots of practice to get ready for some event… It's like they're your team and I'm the coach and we've got to get them to the championship." Mr. A also went on to explain that

he appreciates the OSSLT for bringing focus and purpose to the Applied English curriculum, acknowledging other teachers may not share a similar perspective:

> I think there are probably lots of old-school teachers who say, "Oh that's so stupid; why do we bother with test preparation?" But, I think it's helpful because it tells you what you have to do and you can circle all your lesson plans toward that goal. So I'm not totally against it. I kind of like it. It gives you purpose.

In all three classrooms, test preparation was explicit and included repetitive practice centred on avoiding common OSSLT mistakes. Students were given writing templates and formulas for creating news reports and opinion pieces. While Ms. D tried to "push something in there with everything they did," when students practised writing opinion pieces and news reports, Mr. C focused on "the building blocks of paragraph writing." In addition to topic sentences, this meant "guiding the reader with main ideas" and "building vocabulary." While reviewing a student's duotang, Ms. B discussed everything she and Mr. A had focused on in relation to OSSLT preparation, including multiple uses of quotations, paragraph structure, news reports, and writing templates. She then described their work with students in using exemplars and templates, and the thorough analysis of news reports, focusing on voice (first or second person), tense, and patterns. Students practised using these templates with strong emphasis on topic and concluding sentences. Ms. B believed there were tricks that students needed to be aware of in order to be successful on the OSSLT. It seemed likely that OSSLT items were a consistent focus for these teachers. Writing activities centred on writing newspaper reports, the primary writing task on the OSSLT each year, and sample OSSLT reading passages were used in class to teach students how to prepare for the comprehension portion of the test. Given the potential consequences of failure on the OSSLT, this focus was not unexpected, but it highlights a perspective where the purpose of Applied English was not to teach literacy, or instill an appreciation for the English language, but to prepare students to pass a high-stakes test.

Caring is motivating students. Knowing their students lacked intrinsic motivation to engage in test preparation practices, all four teachers employed their own methods of extrinsic motivation. The most obvious example was Mr. C, who used candy to reward a variety of positive classroom behaviours, such as providing correct answers, punctuality,

participation in class discussions, and completing work in class. However, a more commonly observed strategy was to incorporate students' individual interests and natural curiosities into the lesson. This might include unique stories or anecdotes, or intriguing facts or "tidbits" of knowledge. Ms. D spoke at length about how she liked to incorporate learning and knowledge of different cultures into her teaching because the students "eat it up" and Mr. C spoke about the importance of "setting things up to get students curious." We saw these attempts to incorporate students' interests as a way of supplementing the extrinsic motivators (cakes and candies) with intrinsic motivation.

Caring is the use of ongoing formative assessment. During their interviews, both female teachers discussed the importance of frequent assessment and checking in with students. As Ms. D explained, "To me assessment is, 'How are you doing? How can I help? Are we well here?'" Similarly, Ms. B added, "I just assess 'Do I need more time? Do I need to explain more? Do they need to work on this more?'" While Mr. C did not mention ongoing formative assessment during our interviews, observations of his classroom revealed that he incorporated a minimum of four formative assessments per class. As an example, the introductory activity was always assessed, with immediate feedback given to the students. This was usually accomplished by calling on students to show their answers, or by circulating around the room and checking their responses as they worked. These assessment opportunities (in all three classrooms) were low-stakes and confidence-building. There were no grades associated with the formative assessments, and the information was used solely to inform feedback and further the teaching activities. The intent was to help student be successful on their summative assessments. When it came time for students to complete those summative assessments, the assigned tasks mirrored the formative assessment tasks completed in class. Furthermore, summative tasks were structured to maximize the likelihood of success. Ms. D said that many of her students suffered from test anxiety and so she had students complete "evaluations that take a longer period of time" (e.g., spread over several classes) to minimize the negative impact of a student having a bad day. The teachers in our study linked success on assessments with students' emotional well-being, and believed that student success was critical both to their self-image and their willingness to engage with the course.

Discussion

All four teachers were keenly aware of the challenges facing students in their Applied English class and constructed their teaching practice in ways they felt would best help students overcome these challenges. Key concerns were centered on students' low motivation to do well in school, poor learning skills, and difficulty in passing the OSSLT. Because teachers saw passing the OSSLT as vitally important to students' self-image (and their future), this is where they concentrated their teaching efforts. To a large extent, the teachers in our study enacted Applied English as an exam preparation course. While teachers expressed notions of literacy that emphasized an ability to function in society or in the workplace, the content and manner of instruction was geared toward helping students do well on OSSLT items. Skills which may be reasonably construed to represent functional literacy (e. g., writing emails, analyzing the trustworthiness of online information) were not incorporated into any of the classes we observed.

Noddings (1995) lamented that schools' purpose was "academic adequacy" (p. 366) and this approach was perfectly illustrated by our case study. The teachers' desire for their students to pass the OSSLT was understandable and even laudable, but it had profound impact on how they taught Applied English. The extreme amount of washback is consistent with that predicted using Green's (2007) model. The OSSLT is a test of high importance for all Ontario students, but or Applied English students it is also of high difficulty. Clearly, our four teachers believed that devoting many classroom hours to OSSLT test preparation was in the best interests of the students. The importance of the test for graduation meant that passing the OSSLT took precedence over other desirable course goals, such as developing multi-literacies or an appreciation for literature. Part of the reason for this may have been the fixed mindset the teachers adopted toward their students. All four teachers felt confident in predicting career paths for these students, and these paths were limited to jobs that did not require much reading or writing (e.g., auto mechanic, farmworker, and hospitality worker). It may be that because these teachers did not envision a future for these students where further development of writing skills would be beneficial, they were content to forego many aspects of the curriculum to focus on test preparation.

The extent to which the OSSLT influenced teachers' practice was best exemplified by Mr. A, who saw the OSSLT as providing direction and purpose to the course. The

Ontario English curriculum is very open and flexible, and so teachers are not always sure about the content and strategies to be implemented. A standardized examination like the OSSLT can function as an enabling constraint, giving teachers a concrete goal to attain, and establishing parameters for the course. Washback does not result solely from teachers wanting their students to do well, but also from teachers seeking direction and guidance on how to structure their course. There is ample evidence for this in the literature (e.g., Alderson & Wall, 1993; Berliner, 2011; Saif, 2006; Yeh, 2005).

While we have adopted an appreciative stance toward our participants, we do not wish to ignore the fact that some of their teaching practices are open to criticism. For instance, how are students expected to develop organizational skills or responsibility if they are never required to practise these skills? The extreme levels of structure present in the courses meant these teachers never had to chase a lost textbook or follow up on incomplete homework, but also meant that students never needed to demonstrate initiative or independence. Even though we believe these teachers adopted this approach out of genuine concern for their students, and not for their own convenience, the end result is the same: students faced no negative consequences for having poor organizational skills, nor did they experience opportunities to practise developing these skills in a supportive environment. Likewise, we can rationalize the frequent use of external motivators as an attempt to compensate for the low motivation of these students. However, there are unintended consequences of such teaching practices.

How can we know what actions are in students' best interests, especially over the long term? Will the algorithmic, mechanistic approaches to writing taught to the students in these classes help them pass the OSSLT, or hinder their future development as writers? Will the promise of candy spur a student to focus in Applied English, or will it cause him to lose focus in other courses, because those teachers do not offer any reward for participation? Does requiring students to keep their duotangs in class at all times prevent lost materials and help students maintain a full set of notes, or does it prevent further review and study at home? The answers to these questions are unknown (and likely different for every student), and so teachers must make their own judgements about what actions are in the best interests of their students.

We recognize that positioning teaching to the test as an act of caring is unusual, and perhaps even controversial, but the evidence from our study indicates these four teachers taught to the test because they viewed it as a practice that would benefit their

students. In this respect, it was a profound act of caring. However, this leads to the possibility that for any high-stakes test, caring teachers will teach to the test, narrowing the curriculum, thereby reducing the course to test preparation. While this case study is limited to only four teachers within a single school, the idea that teachers teach to the test because they care deeply about the success of their students is reasonable and aligns with our informal conversations with teachers from all over the world. To determine the extent to which this is true, a large-scale study examining teacher motivations for teaching to the test would be helpful.

Conclusion

Our data revealed seven different ways that the teachers in our case study demonstrated caring toward their Applied English students, and chief among them was teaching to the test. In the context of our case study, washback effects of the OSSLT in the Applied English classroom were prominent because the test was difficult (but passable) for these students. Teachers cared deeply about their students and did not want their students to suffer the negative emotional and scholastic effects of not passing the OSSLT, and so structured much of their teaching as OSSLT preparation and practice. These findings are consistent with Green's (2007) model of washback. Other teacher behaviours, such as the use of rewards to motivate students, and removing students' responsibility for their own learning, were also enacted because these teachers believed these practices were the best way to help students succeed.

References

Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*, *13*(3), 280–297.

Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, *14*(2), 155–129.

Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, *13*(3), 257–279.

Baker, R. M. (2006). *How caring informs teaching* (Publication No. 3239544) [Doctoral dissertation, University of Oklahoma]. https://shareok.org/bitstream/handle/11244/1122/3239544.PDF?sequence=1

Barnes, M. (2017). Washback: Exploring what constitutes "good" teaching practices. *Journal of English for Academic Purposes*, *30*, 1–12.

Barrow, M. (2015). Caring in teaching: A complicated relationship. *Journal of Effective Teaching*, *15*(2), 45–59.

Berliner, D. (2011). Rational responses to high stakes testing: The case of curriculum narrowing and the harm that follows. *Cambridge Journal of Education*, *41*(3), 287–302.

Card, D., & Payne, A. A. (2015). *Understanding the gender gap in postsecondary education participation: The importance of high school choices and outcomes*. Higher Education Quality Council of Ontario.

Csikszentmihalyi, M. (1997). *Finding flow: The psychology of engagement with everyday life*. Basic Books.

Damankesh, M., & Babaii, E. (2015). The washback effect of Iranian high school final examinations on students' test-taking and test-preparation strategies. *Studies in Educational Evaluation*, *45*, 62–69.

Dodds-Urban, V. (1999). Eugene's story: A case for caring. *Educational Leadership*, *56*(6), 69–70.

Education Quality and Accountability Office (EQAO). (2006). *Ontario Secondary School Literacy Test Framework*. https://collections.ola.org/mon/16000/270420.pdf

Education Quality and Accountability Office (EQAO). (2017). Highlights of the Provincial Results: Ontario Secondary School Literacy Test (OSSLT). https://www.eqao.com/en/assessments/results/communication-docs/provincial-report-highlights-osslt-2017.pdf

Gewertz, C. (2019, April 9). *Which states require an exam to graduate?* Education Week. https://www.edweek.org/ew/section/multimedia/states-require-exam-to-graduate.html

Green, A. (2007). *IELTS washback in context: Preparation for academic writing in higher education, Studies in language testing* (Vol. 25). Cambridge University Press and Cambridge ESOL.

Hamlin, D., & Cameron, D. (2015). *Applied or academic: High impact decisions for Ontario students*. People for Education.

International Reading Association. (2014). *Using high-stakes assessments for grade retention and graduation decisions* [Position statement]. http://www.literacyworldwide.org /docs/default-source/where-we-stand/high-stakes-assessments-position-statement.pdf

Kearns, L. (2011). High-stakes standardized testing and marginalized youth: An examination of the impact of those who fail. *Canadian Journal of Education*, *34*(2), 112–130.

Kilickaya, F. (2016). Washback effects of a high-stakes exam on lower secondary school English teachers' practices in the classroom. *Lublin Studies in Modern Languages and Literature*, *40*(1), 116-134.

Kohn, A. (2000). *The case against standardized testing: Raising scores, ruining the schools*. Heineman.

Langer, J. A. (2000). Excellence in English in middle and high school: How teachers' professional lives support student achievement. *American Educational Research Journal*, *37*(2), 397–439.

Lee, G., & Schallert, D. L. (2008). Constructing trust between teacher and students through feedback and revision cycles in an EFL writing classroom. *Written Communication*, *25*(4), 506–537.

Leech, N. L., & Onwuegbuzie, A. J. (2007). An array of qualitative data analysis tools: A call for data analysis triangulation. *School Psychology Quarterly*, *22*(4), 557–584.

Luce-Kapler, R., & Klinger, D. (2005). Uneasy writing: The defining moments of high-stakes literacy testing. *Assessing Writing*, *10*(3), 157–173.

Mendelsohn, D. J. (1989). Testing should reflect teaching. *TESL Canada Journal*, *7*(1), 95–108.

Merriam, S. (1998). *Qualitative research and case study applications in education* (2nd ed.). Jossey-Bass Publishers.

Musoleno, R. R., & White, G. P. (2010). Influences of high-stakes testing on middle school mission and practice. *RMLE Online*, *34*(3), 1–10.

Nelson, G. L. (2000). Warriors with words: Toward a post-Columbine writing curriculum. *English Journal*, *89*(5), 42–46.

Noddings, N. (1984). *Caring, a feminine approach to ethics & moral education*. University of California Press.

Noddings, N. (1988). An ethic of caring and its implications for instructional arrangements. *American Journal of Education*, *96*(2), 215–230.

Noddings, N. (1995). A morally defensible mission for schools in the 21st century. *The Phi Delta Kappan*, *76*(5), 365–368.

Ontario Ministry of Education. (2007). The Ontario Curriculum Grades 9 and 10: English. http://www.edu.gov.on.ca/eng/curriculum/secondary/english910currb.pdf

Popham, W. J. (2001). Teaching to the test. *Educational Leadership*, *58*(6), 16–20.

Saif, S. (2006). Aiming for positive washback: A case study of international teaching assistants. *Language Testing*, *23*(1), 1–34.

Shepard, L. A., & Dougherty, K. C. (1991, April 3–7). *Effects of high-stakes testing on instruction* [Paper presentation]. Annual Meetings of the American Educational Research Association, Chicago, IL, United States. https://files.eric.ed.gov/fulltext/ED337468.pdf

Shih, C. M. (2007). A new washback model of students' learning. *Canadian Modern Language Review*, *64*(1), 135–161.

Spratt, M. (2005). Washback and the classroom: The implications for teaching and learning of studies of washback from exams. *Language Teaching Research*, *9*(1), 5–29.

Stake, R. (1995). *The art of case study research*. Sage Publications.

Stake, R. (2006). *Multiple case study analysis*. The Guilford Press.

Stewart, M. A. (2016). Nurturing caring relationships through five simple rules. *English Journal*, *105*(3), 22–28.

Stobart, G. (2003). Editorial. The Impact of Assessment: Intended and unintended consequences. *Assessment in Education: Principles, Policy & Practice*, *10*(2), 139–140.

Tsagari, D., & Cheng, L. (2016). Washback, impact, and consequences revisited. In E. Shohamy, I. Or, & S. May (Eds.), *Language testing and assessment: Encyclopedia of language and education* (3rd ed., pp. 359–372). Springer.

Uppal, S. (2017). Young men and women without a high school diploma, 1990–2016. https://www150.statcan.gc.ca/n1/en/pub/75-006-x/2017001/article/14824-eng.pdf?st=vmTY3sl2

Volante, L. (2004). Teaching to the test: What every educator and policy-maker should know. *Canadian Journal of Educational Administration and Policy*, *35*. https://files.eric.ed.gov/fulltext/EJ848235.pdf

White, B. (2003). Caring and the teaching of English. *Research in the Teaching of English*, *37*(3), 295–328.

Wong, R. H. K. (1969). Educational effects of examinations on pupils, teachers and society. In J. A. Lauwerys & D. G. Scanlon (Eds.), *Examinations. World yearbook of education* (pp. 360–387). Evans.

Yates, G. C., & Hattie, J. (2013). *Visible learning and the science of how we learn*. Routledge.

Yazan, B. (2015). Three approaches to case study methods in education: Yin, Merriam, and Stake. *The Qualitative Report*, *20*(2), 134–152.

Yeh, S. S. (2005). Limiting the unintended consequences of high-stakes testing. *Education Policy Analysis Archives*, *13*(43). https://doi.org/10.14507/epaa.v13n43.2005

Yin, R. (2018). *Case study research and application: Design and methods* (6th ed.). SAGE Publications.