WALKING IN THEIR SHOES: STUDENTS' PERCEPTIONS OF LARGE-SCALE HIGH-STAKES TESTING

Don A. Klinger
Rebecca Luce-Kapler
Queen's University
Kingston, Ontario

Abstract: With the implementation of the Ontario Secondary School Literacy Test (OSSLT) in 2002, Ontario became the first province in Canada requiring successful completion of a large-scale high-stakes literacy test for high school graduation. We began to explore and analyze students' perceptions of this testing program in terms of their preparation for the test, the test's impact and value, and the potential influence of such a testing program on the students' views about literacy. Our study used qualitative data obtained through focus groups and interviews with students who were either successful or unsuccessful on the OSSLT. The students recalled specific test preparation for the OSSLT, often at the expense of their regular schooling. Following test instructions was a clear message to the students, leading to formulaic test responses and a narrowly expressed view of literacy. Potentially important differences were also found between the successful and unsuccessful students.

Résumé :   En 2002, avec la mise en oeuvre du Test provincial de compétences linguistiques (TPCL), l'Ontario est devenue la première province canadienne à exiger la réussite d'un test à grande échelle et à enjeux importants de littératie afin d'obtenir le Diplôme d'études secondaires. Nous avons commencé l'exploration et l'analyse de la perception des étudiantes et étudiants face à ce test en considérant les éléments suivants : leur préparation au test, la répercussion et la valeur de ce test, et l'influence potentielle d'un tel test sur leur opinion de la littératie. Notre étude est basée sur des données qualitatives obtenues par le biais de groupes de discussion (*focus groups*) et d'entrevues avec des étudiantes et étudiants qui ont réussi ou non le TPCL et rapportent la préparation spécifique au TPCL qu'ils ont subie souvent aux dépens de leur programme scolaire régulier. Les étudiantes et étudiants ont reçu un message clair de suivre les consignes du test les encourageant à donner des réponses selon une formule

Corresponding author: Don A. Klinger, Associate Professor, Faculty of Education, Queen's University, Kingston, ON, K7L 3N6; <don.klinger@queensu.ca>

et à exprimer une opinion étroite de la littératie. Des différences potentiellement importantes furent aussi identifiées entre les étudiantes et les étudiants qui ont réussi et celles et ceux qui n'ont pas réussi le test.

██████████ Amanda is a Grade 12 student in Edmonton, Alberta. Part of her preparation for graduation includes studying for the Grade 12 provincial diploma examinations in her academic courses: English Language Arts 30, Math 30, Biology 30, and Chemistry 30. These examinations will contribute 50% to the final grade that Amanda and other students enrolled in Grade 12 academic courses obtain in these courses. In contrast, Amanda's cousin, Robert, is in Grade 10 in Oshawa, Ontario. There are no provincial Grade 12 examinations for Robert to write in Ontario; however, Robert will have to successfully complete the Ontario Secondary School Literacy Test (OSSLT)[1] before he graduates. His graduation rests on one test.

The history of large-scale testing in Canada reflects the changing pedagogical, political, and social perspectives of each province or territory. While at times there has been an almost complete absence of large-scale testing programs across the country, there also have been eras in which large-scale testing predominated in student evaluation (e.g., Nagy, 2000; Ontario Department of Education, 1968; Putman & Weir, 1925; Raphael, 1993). For example, Amanda's father does not recall writing any such large-scale examinations as a high-school student in British Columbia in the 1970s, but Amanda's grandmother still remembers the pressure she felt as she prepared for the high-school exit examinations that were the sole determiner of her high-school grades.

Currently, public education is again in an era of large-scale testing. However, the current models of large-scale testing now include an increasingly important system accountability component alongside the measurement of student achievement (Earl & Torrance, 2000; Firestone, Mayrowetz, & Fairman, 1998; Ryan, 2002). These testing programs are predicated on the belief that large-scale testing programs can affect changes in policy, curriculum, and practice to better focus instruction and increase subsequent student achievement and school performance (e.g., Delandshere, 2001; Madaus & Kellaghan, 1992; Ontario Royal Commission on Learning, 1994; Ryan, 2002). The most dramatic example is the *No Child Left Behind Act* (2002) in the United States that uses test data not only to reward but also to sanction schools in an attempt to ensure student achievement of fundamental skills (e.g., Wilson, 2006). Currently, every province

and territory in Canada has one or more ongoing large-scale testing programs, albeit largely without the rewards and sanctions of the American models (Crundwell, 2005; DeLuca, Brunette, Miller, & Klinger, 2005).

With the completion of the Royal Commission on Learning (1994), Ontario became one of the most recent provinces to re-enter the realm of systemic large-scale achievement testing of school children[2] (Earl & Torrance, 2000; Volante, 2007; Wilson, 2006). The first sets of examinations were administered to children in Grades 3 and 6 in the spring of 1997, focusing on reading comprehension, writing, and mathematics. The examinations were considered low-stakes because they had no influence on the measurement of individual student achievement. Grade 9 mathematics tests for both the academic and applied programs followed in 1998. Again, the results had no official influence on individual student achievement; however, teachers were encouraged to use a portion of the examination to help determine students' grades.

In the 2001/2002 school year, schools administered the Ontario Secondary School Literacy Test (OSSLT) for the first time. This test represented a new realm in high-stakes, large-scale testing not only in Ontario but also in Canada. With the implementation of the OSSLT, Ontario became the first Canadian province to require successful completion of a large-scale literacy test for high-school graduation.

The purpose of the OSSLT is to determine whether students have acquired the reading and writing skills across the curriculum that they are expected to have learned by the end of Grade 9. Since its introduction, the Education Quality and Accountability Office (EQAO) has made several modifications to the OSSLT. For instance, students who do not successfully complete the OSSLT can now obtain the graduation requirement through successful completion of the Ontario Secondary School Literacy Course (OSSLC). More recently, the EQAO removed the separate reading and writing components—both of which students had to successfully complete—and now provides a single literacy score.

PURPOSE

While such modifications attempted to address concerns and issues regarding the impact of the test on students, there has been little research on the OSSLT to examine how the initial design and imple-

mentation of the OSSLT affected or was perceived by those students writing the OSSLT. With its ties to successful graduation, the OSSLT represents a new form of literacy testing in Canada, and its impacts on students have not been examined. Therefore, the purpose of our study was to begin to explore and analyze the effects of a high-stakes, large-scale testing program on students within an educational context where there was no recent history of such testing. Using the perceptions of students, based on focus groups and interviews, we examined the perceived impacts on, and the experiences of, students who are faced with an examination program that demands success in order to graduate. In order to examine the spectrum of experiences, we included students who had successfully and unsuccessfully completed the OSSLT and also students who were preparing to write the OSSLT for the first time, speaking to both those whom the school expected to be successful and those expected to be unsuccessful. Specifically, our research focused on issues of test preparation at the school, students' approaches to preparing for and writing the OSSLT, and the students' views on the impact and value of the OSSLT. We were also interested in how such a testing program influenced these students' views on literacy and high-stakes, large-scale testing.

In order to study the extent to which these diverse groups of students varied in their practices and perceptions surrounding the OSSLT, we developed the following research questions: (a) What specific preparation programs did the students complete in preparation for the OSSLT? (b) What was the impact of the OSSLT with respect to school, future education, and career directions? (c) What similarities and differences could be found between successful and unsuccessful students' approaches to and perceptions of the OSSLT? These questions were intended to address our overarching concern in this study: to begin to investigate the potential impact of a high-stakes, large-scale testing program on students.

The Ontario Secondary School Literacy Test

Provincial testing programs in Ontario are mandated by the Ontario Ministry of Education. However, the actual design, development, administration, and scoring of provincial testing programs is completed by the EQAO, an "arms-length" educational organization in Ontario. As part of its mandate, the EQAO is responsible for the OSSLT. Two key purposes of the OSSLT are to provide reliable data about the reading and writing skills of students in Grade 10 and to ensure that graduating students in Ontario demonstrate fundamental literacy

skills (EQAO, n.d.c). The OSSLT is designed to measure students' cumulative literacy skills through nine years of schooling across curricula as opposed to merely being a test of English (or French) achievement (EQAO, 2003; EQAO, n.d.b, p. 1).

The OSSLT examines both reading and writing in a print context. At the time of our data collection, the test occurred over two days with two hours provided each day.[3] The reading component contained 12 reading selections designed to measure informational, graphic, and narrative texts. Multiple-choice, short answer, and written explanation test items were used for the reading component. In the writing component, students were required to complete four writing tasks, a summary, a series of paragraphs expressing an opinion, a news report, and an information paragraph. For each task, the development and organization of relevant ideas, writing structure, and conventions of writing were the key aspects to be assessed. Students completing the OSSLT receive notification of their achievement on the test. Unsuccessful students also receive feedback regarding identified weaknesses on the test. The purpose of this feedback is to help support unsuccessful students as they prepare to meet the graduation requirement in a future sitting of the test.

Although there are several examples of high-stakes testing programs in Canada in which the results are used in part to determine grades (see, for example, Alberta Education, n.d.; British Columbia Ministry of Education, n.d.; Nova Scotia Department of Education, n.d.), the OSSLT is the first high-stakes testing program in Canada to make the successful completion of a high-stakes literacy examination a graduation requirement.

Impact of High-Stakes Testing on Teaching and Learning

High-stakes testing is defined in terms of the consequences and importance for educational practices such as marks, grade promotion, graduation, or admission (Santrock, Woloshyn, Gallagher, Di Petta, & Marini, 2004). As more importance is placed on standardized testing programs, the need to analyze the intended and unintended consequences of testing students increases. The argument here is that increasing attention needs to be focused on the effect testing is having on students. Research has consistently found that classroom teaching, student choices, and assessment practices are influenced by high-stakes, large-scale tests. However, the consequences of these influences may or may not be desirable (e.g., Lane & Stone, 2002;

Rogers, 1991). For example, Darling-Hammond, Ancess, and Falk (1994) pointed out that a consequence of high-stakes testing was a narrowing of curriculum, reduced instructional time in favour of test preparation activities, and teaching practices that were more test-like. These concerns continue today.

The impact of large-scale, high-stakes testing falls under the notion of consequential validity (Messick, 1989). Literature contains its share of proponents who, with reference to the use of high-stakes tests, point out positive consequences (e.g., Cizek, 2001; Ward, 2000), and critics who call attention to a plethora of negative consequences (e.g., Brennan, 2001; Eisner, 2001; Jones, 2001; Shulha & Wilson, 2002; Thompson, 2001). While the notion of consequential validity remains an important concept, very little research has been conducted examining students' perceptions about the consequences of such testing programs. In one example, Roderick and Engel (2001) examined low-achieving students in the Chicago Public School system, a board requiring students to pass large-scale tests in reading and mathematics and the Iowa Tests of Basic Skills to be promoted into the next grade. Roderick and Engel were able to classify students based on motivation and willingness to work hard to prepare for the tests. Among their conclusions was the caution that one should not assume that low-achieving students would respond negatively to new policies that emphasize achievement through high-stakes testing. The continuing diversity of findings points to the need to better understand the perceptions of both successful and unsuccessful students with respect to high-stakes testing.

METHOD

We collected the data for the study from students in two regions in Ontario, a smaller urban centre (Kingston) and a large urban centre (Toronto). We chose schools that were representative of the surrounding community but were not populated exclusively by extremely high- or low-achieving students (based on EQAO Grade 9 and Grade 10 results). The Kingston school had a slightly higher than average proportion of students pass the OSSLT, whereas the school in Toronto had a lower than average proportion of students successfully complete the OSSLT. Following established guidelines, two separate procedures were used to obtain students' perceptions regarding the OSSLT: individual interviews and focus groups (Krueger & Casey, 2000; Patton, 2002). In the first, individual interviews were conducted with 16 students who were preparing to write the OSSLT for the first time.

We audiotaped the 30-minute interviews during non-instructional times, lunch hour, after school, or non-class periods. On the basis of information provided by school counsellors about students' current literacy achievement and ability to articulate their views, we selected an equal number of male and female students from the two different centres and classified them as either "Likely to be successful" or "Unlikely to be successful." Sampling continued until we had interviewed an equal number of "likely" and "unlikely" to be successful students in each of the schools. Interviews avoided potential stigmatization of students who were preparing to write the OSSLT while allowing us to obtain specific details along with personal beliefs and feelings regarding their thoughts and perceptions. The interviews consisted of five structured open-response questions, enabling the students to describe their feelings and expectations of the OSSLT. For example, the third question was "What impact do you believe the test results will have on your (a) feelings about school? (b) educational program? (c) career and further schooling plans? and (d) opinion of the test?" Other interview questions focused on the preparatory activities occurring in the schools and amongst students, students' knowledge of the test, and students' stress related to the upcoming examination (see Table 1). Where appropriate, short, probing questions were used to obtain more detailed responses. For consistency purposes, the interviews were conducted by two research assistants with interview experience.

Our second procedure involved four focus groups with groups of five to eight students who had completed the OSSLT and had received their results. The focus groups were moderated by research assistants along with the principal researchers. Working with the school counsellors, we used purposive sampling to identify and choose students who would best articulate their personal perceptions while maintaining a gender balance. In each school, one focus group consisted of students who had been successful and one of students who had been unsuccessful on the OSSLT. The focus groups occurred after school, lasted approximately 60 minutes, and were audiotaped. The focus groups were similarly structured in terms of questions, focusing on school and student preparation for the OSSLT, the OSSLT process, students' perceptions and attitudes regarding the OSSLT, and the impact of the test on future educational and professional opportunities. However, due to the nature of focus groups, the questions were not as structured as in the interviews, enabling us to allow the ongoing conversations to better explore themes and ideas. Probing or directional questions were used to maintain the conversational focus or explore a new theme when the conversation seemed to reach a natural end. We also worked to ensure that each of the participants was part of the ongo-

ing conversation. In total, data were obtained from 42 students, 22 classified as successful or likely to be successful and 20 classified as unsuccessful or unlikely to be successful, containing approximately 12 hours of transcripts.

The research assistants transcribed the audiotaped focus groups and interviews verbatim. We first read the transcripts together to identify broad themes and understand the general perceptions and views of the students. This initial coding used the questions to help identify the major themes, test impact, pressure, and test knowledge. We then independently gave the transcripts a second detailed reading to code individual statements into more detailed themes and constructs. Beginning with one of the focus groups, we identified and discussed the constructs found independently in the data. Through discussion, these were clarified and agreed upon for use in

Table 1
Interview and Focus Group Questions

Interview Questions
1. Are you or have you been involved in any special preparation for the writing of the literacy test? If so can you describe these?
2. Please describe your feelings and expectations as you prepare for the test.
3. What do you know of the format and content of the test?
4. What impact do you believe the test results will have on your:
    a. Feelings about school?
    b. Educational program (e.g., future course selection)?
    c. Career and further schooling plans?
    d. Opinion of the test?
5. Think forward to the day of the test. What do you expect the experience to be like?

Focus Group Questions
1. In terms of preparing for the Grade 10 literacy test did you do anything to prepare for it? Personally or in your classes.
2. Tell me about the conditions under which you wrote the test? Tell me about that day.
3. Can you talk about the test specifically? How the test was structured?
4. How did you feel when you were writing the test?
5. What kind of strategies did you use to finish the test?
6. Do you think this test is going to affect:
    a. What courses you are going to choose for Grade 11?
    b. Your career choices?
    c. Other decisions?
7. Do you think it was a good test of literacy? Did it capture literacy?

subsequent transcripts. Any new constructs identified in the transcripts were discussed before being added to the detailed list. A set of 20 constructs provided sufficient coding to represent the data in the transcripts (see Table 2). As an example, under the general theme of test impact, specific constructs included *test impact: work*, *test impact: education*, *future course selection*, and *future*. We used Excel (2004) and then *Concordance* (Watt, 2004) to organize the coded statements. Using these software programs, we were able to code the data and sort the students' statements according to identified themes and constructs, while maintaining the sources of the statements. Due to the qualitative nature of the study, we purposefully did not tally the number of interviewees that expressed a specific sentiment. Nor did we quantify the proportion of students who made or agreed with specific statements in the focus groups. Rather we used the sorted students' statements to identify commonly expressed perceptions and points of view across individuals and focus group participants.

Table 2
Identified Themes and Codes

Confidence	Teacher pressure
Difficulty	Test alternatives
Fairness	Test content
Future	Test impact: education
Future course selection	Test impact: work
Literacy definition	Test impact: worry
Pressure	Test knowledge
School preparation	Test process
Self awareness	Test reporting
Self preparation	Test value

We defined a perception or viewpoint as common if it was expressed by at least half of the interviewees in a particular school and at least half of the participants within two or more of the focus groups expressed or agreed with the viewpoint. For the purposes of comparisons across the groups (successful or unsuccessful) or schools, we considered a difference to exist if twice as many of the participants in one group or school expressed a point of view as compared to the other, or if the viewpoint was not expressed by participants in one group or school but was commonly expressed by participants in the other group or school. We also included and noted less common views and thoughts if they provided valuable insights into students' knowledge, understanding, and perceptions.

For the current study, we were particularly interested in the students' perceptions regarding the testing process, the purposes of the test, and its connection to schooling and future opportunities. As might be expected, much of the data revealed student conceptions (or misconceptions) regarding the relevance, importance, and impact of the test to them as individuals. We also detailed some of the test preparation activities that were occurring, albeit from a student's perspective.

RESULTS

We present the results of the interviews and focus groups to reflect the testing timelines beginning with test preparation, through the testing process, and the resulting perceived impact of the test on the sample of students we spoke to in these two schools. Our intent was to best describe the views and perceptions of these students rather than to make generalized statements regarding the OSSLT or students as a whole. The richest data tended to come from the students who were either successful or deemed likely to be successful. These students responded more openly and insightfully about their views and required little prodding to expand on their thoughts and ideas. Students classified as unsuccessful or unlikely to be successful were somewhat more reserved in their responses, although they would expand on their answers with prodding or encouragement during the interviews. The focus groups seemed to provide a more comfortable atmosphere for the unsuccessful students to recall and describe their experiences, as these students were more forthcoming with their views than the unlikely to be successful students we interviewed. Interview and focus group transcripts were coded based on the school location (1 or 2), the data collection procedure (**In**terview or **F**ocus **G**roup), and the student group (**S**uccessful or **U**nsuccessful). Hence a direct quote from an unsuccessful student participating in the focus group at the first school was coded 1FGU. We provide direct quotes to provide section headings, illustrate the commonly expressed views of students, or to present unique but potentially important or insightful perspectives. Pseudonyms were used in the transcripts and the use of direct quotes in order to fully protect the identity of the students.

Just Follow the Instructions

Given that the OSSLT represented a new form of high-stakes testing for students, it is not surprising that teachers and the school administration would develop policies and procedures to help prepare

students for the test. Throughout the interviews and each of the focus groups, the students spoke about practice activities or worksheets to help them prepare for the OSSLT. The descriptions of many of these activities paralleled the OSSLT tasks and items. However, it is unclear the extent to which these activities benefited the students in terms of developing the foundational literacy skills to be demonstrated on the OSSLT. Stephanie, one of the students who successfully completed the OSSLT, concluded the act of test preparation actually defeated the purpose of the OSSLT: "I don't think the preparation was necessary because to truly test someone's literacy, I don't think that they should be able to prepare and have all these practice tests" (1FGS). Nevertheless, based on the students' comments, the teachers in the schools were working to prepare students for the OSSLT, both technically and emotionally. Some teachers would try to relax students; others would stress the need for students to be successful. Parma reacted to these teachers rather uniquely, recalling the teachers were "freaking out and all the students are just like 'stop, it's not that big a deal'" (1InU).

Test preparation activities in the two schools included early identification of students who might be at risk for failure on the test. Students completed practice tests in Grade 9, and these results were used to provide further literacy practice to struggling students. For example, some struggling students identified attending after-school literacy sessions or a literacy camp to help them prepare for the literacy test. In the time before the test and depending on the school or class, students completed practice reading and writing tasks or worksheets during their English classes, during dedicated literacy times in other classes, or during TAP (biweekly Teacher Advisor Program class meetings). Different recollections surfaced between the interviews and focus groups within each school, suggesting these procedures were either inconsistent or were changing over time.

While the use of various classes to support preparation may have been an attempt to demonstrate that literacy is cross-curricular, the students struggled to see the value of these preparatory activities or the reasons they were occurring in these other classes. Students in both schools, and in particular the unsuccessful and unlikely to be successful students, did not link the test preparation activities to literacy development, but rather saw them as a distinct set of activities purely designed to help students successfully complete the test. Melissa highlighted this lack of connection while also illustrating the struggle these students had in defining literacy: "They may just

study hard for that but then you're at five years down the road they'll totally forget everything about literacy" (2InU). This struggle is further exemplified through the commonly stated belief of successful and unsuccessful students throughout the focus groups and interviews that literacy training was the responsibility of the English teacher or was simply a measure of reading and writing. Time devoted to literacy development in other subject areas was out of place, and, as one student stated, the teachers that are not English teachers "are not even qualified to teach these certain things" (1FGS).

As the date of the OSSLT approached, students sought out further information and supports to help them prepare for the test; however, and not surprisingly, the degree to which these students completed such activities varied. More surprisingly, perhaps, those students at the greatest risk for failure were also the least likely to report they engaged in self-preparatory literacy activities. Only three of the unlikely to be successful students we interviewed noted independent preparation for the test, and the large majority of students in both of the unsuccessful focus groups stated they did not do or recall any self-preparation. In contrast, over half of the successful students in both the interviews and focus groups participated in some form of self-preparation, and these students were also able to describe this preparation in detail. Overall, we observed that the successful students were more aware of the OSSLT and were generally confident about their success on the test. Successful students described their efforts to gather evidence about the OSSLT through learning about the structure and expectations of the test, completing preparatory activities, or obtaining teacher feedback on their progress. While the unsuccessful students were also confident about their success on the test, their reported preparation for the test was never expressed in the same detail or to the same extent as the successful students. Interestingly, although few students expressed doubts about being successful on the OSSLT, when doubts were expressed they more commonly came from successful or likely to be successful students.

The differences in preparatory activities also resulted in a qualitative difference in the students' knowledge of the OSSLT. All of the students had an overall knowledge of the format and structure of the test. Both the unlikely to be successful and the likely to be successful students generally knew about the two main writing samples, the opinion piece and the newspaper article, although two of the unlikely to be successful students did not believe there would be any writing required. More importantly, the likely to be successful students pre-

paring to write the OSSLT demonstrated a much deeper knowledge of the test expectations. For example, these students could describe the structure of the writing tasks to develop and defend an opinion or discuss issues of reading comprehension. This qualitative difference in knowledge was also observed in the students who completed the OSSLT. The unsuccessful students in both focus groups could recall visual images from the test and of their struggles to comprehend the meaning of the image. In contrast, the successful students were more likely to recall specific test questions. In both of these focus groups, the recollection of a specific item led to a prolonged discussion regarding specific items or experiences with the test items. Anna provides an example of this detailed memory. In thinking back, she described one particular question that frustrated her: "Especially the hyphen one—I was pretty mad about that" (1FGS).

The consistent message the students reported receiving from the teachers and the preparatory activities was the importance of following the test instructions. The successful students in the focus groups and approximately half of those interviewed recalled teachers reminding them to read the test items carefully and follow the instructions because deviation from the rules could result in failure. Teachers' warnings included "if they say to write three paragraphs you have to write three paragraphs exactly. So if you write an extra paragraph you fail that section" (2InS). A second example is from Ryan, who believed that after completing the test, "I was pretty sure I was going to pass but if say I forget to indent a paragraph where they told me to indent it and they failed me, which was one of the rules, then you have to do it again" (1FGS). Students thought such strict rules were unfair but believed them to be true since as one student put it, "that's what our teachers were telling us" (2InS). While such strict rules were not actually in place, such exaggerated beliefs are further evidence of student confusion about the definition of literacy and the purpose of the OSSLT, since such strict rules do not examine literacy skills but simply "test how well you follow instructions" (2FGS).

They Just Gave the Test and Said "Do"

The administration of the OSSLT is controlled by the school administration using procedures outlined by the EQAO. While there were clear instructions to ensure standard administrative procedures over the two-day testing period, variations occurred not only in the locations in which the test was written but also in the interactions with the teacher invigilators. The cafeteria, the gymnasium, or even

designated classrooms were used to administer the test. Hence it is not surprising the experiences of students who completed the test differed not only across the two schools but also within each school. As mentioned above, students received the message regarding the importance of strictly following the test instructions. A few students recalled a sense of formality to the test instructions given by the teacher invigilators. After receiving instructions from the teacher invigilator, Omar recalled, "I think the thing that I was most worried about was ripping my sticker. They gave us all stickers that we had to shut it with, and it said if this sticker is ripped this test will not be marked. So we're all like trying to put it on" (1FGS). Other than providing the initial instructions, the teachers seemed to avoid any other interactions with students. According to Stephanie, one teacher wished the students good luck and then commented, "Oops, I can't say that" (1FGS). Students' comments indicated this lack of interaction, including the inability to ask for clarification, was both discouraging and unnecessary.

I Think It's More of a Formula

All but one of the successful students in the higher achieving Kingston school and half of those from the Toronto school described approaches to writing the OSSLT that substantially differed from those described by the unsuccessful students. The successful students commonly either read through the examination and identified items they could easily answer to bolster their confidence or first completed the longer items having a higher test value. These students also knew such items would be near the back of the test booklet. Michael recalled, "The first day I just did it front to back and then I found that I was rushing and it wasn't my best work so then I found that doing the questions toward the back of the book, the bigger questions, I could finish those and do well on them" (1FGS). Anna quickly echoed this strategy: "I would do the ones that were big because I thought they were probably most important and then the short answer ones, I did those last" (1FGS). In the second school Julie used a similar strategy as she remembered, "I looked through it so then I managed my time first and I saw which was hardest and I started with that one" (2FGS). In contrast, with one stated exception, all of the unsuccessful students reported they simply went from the front to the back of the test books, doing those items they believed were easier—the short answer and multiple-choice—or as Riley described, he "did the stuff that wasn't going to take as long as the other stuff just to make sure that I'd have enough time for the longer stuff" (1FGU).

These differences in test-taking strategies also extended to the methods used to answer the questions. The successful students recalled specific strategies to complete the questions. As Petra summarized, "sometimes I was thinking as I was writing I was like what are they trying to like teach me, what are they trying to evaluate with asking these questions?" (1FGS). Although such strategies likely provided avenues for success, they also resulted in recognition of the test's limitations. Ultimately, the students we spoke to concluded the test focused more on demonstrating formulaic writing structures as opposed to literacy. "If [students] put the right things in the right place they'd pass. You really didn't have to be creative or imaginative at all" (1FGS). As with other aspects of the OSSLT, this formulaic approach to testing literacy skills creates difficulties in students' perceptions of literacy. As Omar concluded, "So now being literate is like knowing how to write a certain way and read a certain way and think a certain way" (1FGS).

It Should Be on How You Write and Comprehend, Not on How Fast You Can Do It

While concerns about reading difficulty, testing and marking procedures, and test anxiety were mentioned in the focus groups, the most commonly voiced concern by the unsuccessful students at both schools was the lack of time to complete the test. Jason recalled, "None of us had enough time" (2FGU). While Jason was exaggerating, others also noted not having time to complete the test and observing other students who were writing until time expired. Kyle and Lisa blamed this on the amount of reading required to complete the test, stating, "There's too much reading in there" (2FGU) and "They give us a story that's two pages long" (2FGU). Tyler, who admitted he was a slow reader, believed this amount of reading caused his failure because he "didn't get all the reading done" (1FGU). The front-to-back completion strategy of the unsuccessful students may also have proved to be problematic given the time constraint. Lauren recollected, "I didn't finish either [day] because I ran out of time. There was a big paragraph that you had to write at the end of each one and I missed both of them" (1FGU).

Students preparing to write the test similarly voiced their concerns about having sufficient time. Interestingly, this concern was expressed by students in both the likely and unlikely to be successful groups. While students could be identified prior to the administration of the OSSLT to receive extra time to complete the test, only one of

the interviewed unlikely to be successful students noted this had oc-
curred for her. This accommodation was a surprise to students such
as Rina, who said she was not even aware such provisions existed.

It's Just Kind of Like an Obstacle and You Have to Finish It

When we asked about the impact and value of the OSSLT, the stu-
dents expressed a multitude of opinions. While Parma may have
believed the test was not that big a deal, the majority of the students
we spoke with generally commented that they were worried and
nervous about the test. These students recognized the OSSLT was a
high-school requirement. Interestingly, while recognizing the impor-
tance of the OSSLT for graduation, few students believed it would
be a real barrier to their future decisions, largely because they could
repeat the OSSLT. Rather, the impact of failure on the OSSLT was
recognition that further literacy training would be required prior to
graduation. Similarly, the students we interviewed did not generally
believe the OSSLT would impact their future career choices. However,
one unique but interesting comment came from Justin, who "want[s]
to be a mechanical engineer or a DJ [disc jockey]" (2InU). He felt that
failure on the OSSLT would help him decide his career choice since,
as he concluded, failure would still allow him to be a DJ. What we did
observe was a fundamental difference in the career aspirations of the
successful and likely to be successful as compared to the unsuccessful
and unlikely to be successful students; the successful students gener-
ally spoke of post-secondary education as a given. For these students,
the only potential impact of the OSSLT, if any, would be the location
and direction of their post-secondary studies. Unsuccessful students
tended to view post-secondary education as a potential option that
success on the OSSLT might make easier.

Is the OSSLT valuable and important to students? Our interviews
uncovered a wide variety of responses both for and against the OS-
SLT while also providing alternatives and suggestions for improving
the OSSLT. For successful students such as Stephanie and Anna, the
lack of feedback limited the usefulness of the test.[4] Anna highlighted
this frustration, concluding, "We are getting absolutely no informa-
tion so how can we improve on our skills?" (1FGS). In spite of this
lack of useful information, successful students were more likely than
the unsuccessful students to see a potential value to the OSSLT. As
Marianna stated, "I think [the test is] good, it gives the government,
and gives the parents, and gives everybody an idea on where you
stand" (1InS). Andrew thought that successful completion of the

OSSLT would be good preparation for students, so "in the long run it [the OSSLT] will really help them" (1InS).

Certainly, support for the test was in the minority amongst the successful and likely to be successful students, and the unsuccessful and unlikely to be successful students were almost unanimously against the OSSLT. A common sentiment was that the test was "just a waste of time and money" (1InU). However, a second common complaint from students in both schools was that the test was duplicating the work from their English classes. Two specific comments from the focus groups and interviews illustrate this belief: "we're doing the same thing on the Literacy Test that we do in our other English classes and we're passing that. So what's the point of this?" (2FGU) and "What's the point of having two end of the year exams for English?" (2InU). Once again, these comments reflect a point of view we commonly found across these students that tended to view the OSSLT as a test of the English curriculum rather than the cross-curricular literacy espoused by the EQAO.

DISCUSSION

Our research used qualitative methodologies to begin to examine the perceived impact of the OSSLT on students. Our questions focused on test preparation, impact, and value. Interviews and focus groups enabled us to explore some students' perceptions and knowledge of the OSSLT while also making preliminary but important comparisons between those students who were either successful or likely to be successful with unsuccessful or unlikely to be successful students writing the OSSLT. According to these students, their schools provided specific test preparation for the OSSLT, often in the year prior to the OSSLT. Not surprisingly, the preparatory activities varied in nature and in delivery. Regardless, students generally reported that these test preparation activities interfered with their schooling. The preparatory message the students recalled was the importance of close adherence to the instructions in the OSSLT. Any deviation from the strict instructions could result in failing scores. Some of the students used this information to criticize the OSSLT as being a limited test of literacy since it seemed to measure the ability to follow strict guidelines rather than literacy skills. While the actual instructions for the OSSLT do remind students to follow the instructions and use appropriate forms of writing, specific details about length or style are not given. Instead, students are instructed to ensure that paragraphs are clearly divided and that the number of lines provides an approximate

guide to the length of the expected answers. Such differing messages do not help students preparing for and writing the OSSLT. Nor do such contradictions instill confidence in the OSSLT itself. Thus it will be important to determine if other students in Ontario express the perceptions we found in the students we interviewed.

Our conversations with students enabled us to identify important differences amongst the groups of students we included. Of particular interest were the qualitative differences in the overall knowledge of the test and the manner the successful and unsuccessful students wrote the OSSLT. Although to a lesser extent, these differences also seemed to exist between the higher and lower achieving schools. Successful students demonstrated a deeper understanding of the test. These students provided well-defined strategies while writing the OSSLT that they also adapted as necessary. In contrast, the unsuccessful students expressed only a general knowledge of the OSSLT and used much simpler and less strategic approaches while writing the test. While limited, our preliminary findings identify a potentially important difference between students who are successful or unsuccessful on the OSSLT, or between higher and lower achieving students. These differences could also help explain the unsuccessful students' higher reported incompletion rates on the relatively important OSSLT items at the end of the test booklets. Recent changes to the OSSLT have reduced the time problem. Nonetheless, our results illustrate a potentially important underlying skill difference that would continue to hamper struggling students. None of the students we spoke to described any form of preparation related to this skill difference. Based on the 2007 results, almost one third of students enrolled in applied education programs failed the OSSLT, and 50% of students retaking the OSSLT failed on their most recent attempt (EQAO, n.d.b). Is it the case these students need direct instruction on assessment strategies and organization? Our findings provide a direction for future research and potential supports for low achieving students who continue to struggle on the OSSLT and with schooling in general.

Although not as dramatic as the difference found between the successful and unsuccessful students, the students from the Kingston school tended to have more information about the OSSLT. Generally, the demographic profile of this school was higher in terms of achievement and neighbourhood wealth as compared to the provincial average and the other school in our study. It also had a much higher success rate on the previous administration of the OSSLT relative to the provincial average and the second school in the study. Future

research will enable us to determine if similar findings occur in the larger Ontario population of high-school students.

Finally, students provided mixed views regarding the value of the test, often citing the same reasons used in the ongoing academic and professional debates about large-scale testing programs. In support of the OSSLT, students mentioned the need for consistency, school monitoring, ensuring academic standards, or identifying students in need. In contrast, those against the OSSLT mentioned the limited value of the test, duplication of information, or lack of fairness, to name a few. While we did not determine the foundations and sources for these opinions, it would seem the OSSLT was certainly a topic for discussion throughout both schools both inside and outside the classrooms.

Throughout our discussions with students, we were continually amazed at the notions of literacy espoused by the students. Students viewed the OSSLT as an English Language Arts test. None of the students considered the test to have implications for their education in other subject areas. The test seemed to impede students' understanding of literacy and even the importance of literacy. As Robert said to Amanda on the phone, "We already know how to read and write. What do we need the literacy test for?"[5]

CONCLUSION

The mandate of the EQAO is "to ensure greater accountability and contribute to the enhancement of the quality of education in Ontario" (EQAO, n.d.a). As part of this mandate, the EQAO "values the well-being of learners above all other interests" and "values only that information which has the potential to bring about constructive change and improvement" (EQAO, n.d.a). Certainly, the EQAO works to deliver a consistent message to the education stakeholders in Ontario. This effort includes various information links and supports on the EQAO website, a field team of seconded educators working with schools and boards, the EQAO annual report, and, most recently, symposiums and conferences for Ontario educators (e.g., *What EQAO Assessments Tell Us About Student Learning*; *Large-Scale Assessment: Supporting the Everyday Work of Schools*). Nevertheless, our research with small but divergent groups of high-school students suggests the EQAO mandate and values may not be either clearly communicated or understood by the most important group of stakeholders, the students who write the OSSLT. Even less clear is the extent of these issues and any ongoing

barriers to clear communication and understanding. All of the students we spoke with reported that teachers provided opportunities to prepare for the OSSLT, but they also noted informal comments from teachers that may have affected either the students' views of the OSSLT or the manner in which students approached the test. Similarly, students influence each other. If the students' preconceived beliefs and misconceptions we found are prevalent throughout Ontario, the EQAO will continue to struggle to clearly communicate its message, and the OSSLT will fail in meeting the mandate and values of the EQAO. The EQAO is responsible for ensuring its message is clearly understood. The EQAO must also revisit its mandate and values to ensure there is a fit between its own values and the testing programs it uses to support these values.

ACKNOWLEDGEMENTS

NOTES

1. The Ontario Secondary School Literacy Course (OSSLC) was implemented with the 2003–2004 school year. Students who have previously failed the OSSLT or have deferred writing the OSSLT may enrol in the OSSLC. Successful completion of the OSSLC fulfills the literacy graduation requirement. The OSSLC was not an option for students at the time of this study.

2. During the 1993/1994 and 1994/1995 school years, the Ministry of Education in Ontario did administer reading and writing tests for all students in Grade 9.

3. Students who have an Individualized Education Plan (IEP) identifying they require more time to complete assessments may have up to four hours to complete the OSSLT.

4. Unsuccessful students receive feedback about their performance designed to help their preparation for the next administration. In contrast, successful students only receive notification that they have successfully completed the OSSLT.

5. This quote is a compilation from statements made by three of the interviewed students. Hence the wording has been modified to reflect the meaning the three students were trying to convey.

REFERENCES

Alberta Education, Ministry of Education. (n.d.). *Students. Programs and services*. Retrieved March 15, 2007, from <http://ednet.edc.gov.ab.ca/students/>.

Brennan, R. (2001). Some problems, pitfalls, and paradoxes in educational measurement. *Educational Measurement: Issues and Practice, 20*(4), 6–18.

British Columbia Ministry of Education. (n.d.). *Provincial examinations*. Retrieved February 20, 2007, from <http://www.bced.gov.bc.ca/exams/>.

Cizek, G. (2001). More unintended consequences of high-stakes testing. *Educational Measurement: Issues and Practice, 20*(4), 19–27.

Crundwell, R. M. (2005). Alternative strategies for large scale student assessment in Canada: Is value-added assessment one possible answer? *Canadian Journal of Educational Administration and Policy, 41*, 1–21. Retrieved February 11, 2007, from <http://www.umanitoba.ca/publications/cjeap/pdf%20files/crundwell.pdf>.

Darling-Hammond, L., Ancess, J., & Falk, B. (1994). *Authentic assessment in action: Studies of schools and students at work*. New York: Teachers College Press.

Delandshere, G. (2001). Implicit theories, unexamined assumptions and the status quo of educational assessment. *Assessment in Education, 8*, 113–133.

DeLuca, C., Brunette, J., Miller, T., & Klinger, D.A. (2005, May). *Assessment cultures in Canadian education: The provincial perspectives*. Paper presented at the annual meeting of the Canadian Society for the Study of Education, London, Ontario.

Earl, L., & Torrance, N. (2000). Embedding accountability and improvement into large-scale assessment: What difference does it make? *Peabody Journal of Education, 75*(4), 114–141.

Education Quality and Accountability Office. (2003). *OSSLT curriculum connections*. Retrieved July 20, 2007, from <https://ospace.scholarsportal. info/bitstream/1873/4930/1/10309431.pdf>.

Education Quality and Accountability Office. (n.d.a). *Highlights of the Provincial Results: Ontario Secondary School Literacy Test 2006–2007*. Retrieved July 28th, 2007, from <http://www.eqao.com/pdf_e/07/ 07P018e.pdf>.

Education Quality and Accountability Office. (n.d.b). *Student resources: The Ontario Secondary School Literacy Test*. Retrieved August 20, 2007, from <http://www.eqao.com/Students/students.aspx?status= logout&Lang=E>.

Education Quality and Accountability Office. (n.d.c). *Successful strategies for improving literacy*. Retrieved July 20, 2007, from <http://www. eqao.com/pdf_e/04/04P011e.pdf>.

Eisner, E. (2001). What does it mean to say a school is doing well? *Phi Delta Kappan, 82*, 367–372.

Firestone, W., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis, 20*(2), 95–113.

Jones, L. (2001). Assessing achievement versus high-stakes testing: A crucial contrast. *Educational Assessment, 7*, 21–28.

Krueger, R.A., & Casey, M.A. (2000). *Focus groups: A practical guide for applied research* (3rd ed.). Thousand Oaks, CA: Sage.

Lane, S., & Stone, C.A. (2002). Strategies for examining the consequences of assessment and accountability programs. *Educational Measurement: Issues and Practices, 21*(1), 23–30.

Madaus, G.F., & Kellaghan, T. (1992). Curriculum evaluation and assessment. In P.W. Jackson (Ed.), *Handbook of research on curriculum* (pp. 119–154). New York: Maxwell Macmillan.

Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York: American Council on Education and Macmillan.

Nagy, P. (2000). The three roles of assessment: Gatekeeping, accountability, and instructional diagnosis. *Canadian Journal of Education, 25*, 262–279.

No Child Left Behind Act. (2002). Pub. L. No. 107–10.

Nova Scotia Department of Education: Evaluation Services (n.d.). *Nova Scotia Evaluation Services Division*. Retrieved March 25, 2007, from <http://plans.ednet.ns.ca/>.

Ontario Department of Education. (1968). *Living and learning: The report of the provincial committee on aims and objectives of education in the schools of Ontario*. Toronto: Author.

Ontario Royal Commission on Learning. (1994). *For the love of learning: Report of the Ontario Royal Commission on Learning*. Toronto: Author.

Patton, M.Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Putman, J.H., & Weir, G.M. (1925). *Survey of the school system*. Victoria, BC: Charles F. Banfield.

Raphael, D. (1993). Accountability and educational philosophy: Paradigms and conflict in Ontario education. *Canadian Journal of Education, 18*, 29–45.

Roderick, M., & Engel, M. (2001). The grasshopper and the ant: Motivational responses of low-achieving students to high-stakes testing. *Educational Evaluation and Policy Analysis, 23*, 197–227.

Rogers, W.T. (1991). Educational measurement in Canada: Evolution or extinction? *Alberta Journal of Educational Research, 37*, 179–192.

Ryan, K. (2002). Assessment validation in the context of high-stakes assessment. *Educational Measurement: Issues and Practice, 21*(1), 7–15.

Santrock, J.W., Woloshyn, V.E., Gallagher, T.L., Di Petta, T., & Marini, Z.A. (2004). *Educational psychology* (1st Canadian ed.). Toronto: McGraw-Hill Ryerson.

Shulha, L.M., & Wilson, R.J. (2002, January-February). *Rethinking large scale assessment*. An invited paper presented at a Symposium on Large-Scale Assessment Provincial Testing in Canadian Schools: Research, Policy & Practice, Victoria, BC.

Thompson, S. (2001). The authentic standards movement and its evil twin. *Phi Delta Kappan, 82,* 358–362.

Volante, L. (2007). Educational quality and accountability in Ontario: Past, present, and future. *Canadian Journal of Educational Administration and Policy, 58,* 1–21. Retrieved March 11, 2007, from <http://www.umanitoba.ca/publications/cjeap/pdf%20files/volante.pdf>.

Ward, C. (2000). GI Forum v. Texas Education Agency: Implications for state assessment programs. *Applied Measurement in Education, 13,* 419–426.

Wilson, R.J. (2006, June). *EQAO: Then and now*. Plenary address for the What EQAO Assessments Tell Us About Student Learning Symposium, Toronto, ON.

Don Klinger is an associate professor in Assessment and Evaluation in the Faculty of Education, Queen's University. His research interests include research methods, the examination of psychometric and policy issues of large-scale assessments, program evaluation, and measures of school effectiveness.

Rebecca Luce-Kapler is a professor of Language and Literacy in the Faculty of Education, Queen's University. Her research focuses on writing processes and technologies. Her book *Writing With, Through and Beyond the Text: An Ecology of Language* brings together her work with women writers and her understanding of learning, writing, and teaching. She has been a fiction writer and poet for over 25 years and is the author of a collection of poetry, *The Gardens Where She Dreams*.