

What has experience got to do with it? An exploration of L1 and L2 Test takers' perceptions of test performance and alignment to classroom literacy activities

Christine Doe & Liying Cheng
Queen's University

Janna Fox
Carleton University

Don Klinger
Queen's University

Ying Zheng
Pearson Language Tests

Abstract

The importance of first language (L1) and second language (L2) test takers' experience with large-scale literacy testing has been well documented in educational research. Our study focused on the Ontario Secondary School Literacy Test (OSSLT), a cross-curricular literacy test that is one of the graduation requirements for Ontario high school students. We drew on qualitative data obtained from open-ended questions on a large-scale survey administered to OSSLT test takers. Whether the test takers were preparing to take the test or had already taken the test seemed to play a critical role in L1 and L2 test takers' perceptions. The most salient results highlight the role that test experience had on test takers' perceptions of their test performance and the alignment between the test and their classroom literacy activities.

Author's Note

Correspondence should be addressed to Christine Doe, Faculty of Education, Duncan McArthur Hall, 511 Union Street, Queen's University, Kingston, ON, Canada, K7M 5R7.
E-mail: christine.doe@queensu.ca

Résumé

L'importance de l'expérience des élèves de langue maternelle (L1) et de langue seconde (L2) soumis à un test, dans le cadre de tests à grande échelle en littératie, a été bien documentée dans la recherche en éducation. Notre étude s'est portée sur le Test provincial de compétences linguistiques (TPCL), une épreuve qui doit être réussie pour obtenir le diplôme d'étude secondaire de l'Ontario. Nous avons puisé des données qualitatives obtenues à partir de questions ouvertes d'une enquête à grande échelle destinée à des élèves passant le TPCL. Le fait que ces élèves aient été préparés ou non à passer le test ou s'ils l'avaient déjà passé, s'est révélé jouer un rôle crucial dans la perception des élèves de langue maternelle (L1) et de langue seconde (L2). Les résultats les plus marquants ont mis en évidence l'importance de l'expérience des élèves avec les tests sur la perception de leur performance et, de la corrélation de ce test avec les activités de littératie en classe.

Note de l'auteur

Toute correspondance doit être adressée à Christine Doe, Faculté d'éducation, Duncan McArthur Hall, 511, Union Street, Queen's University, Kingston, ON, Canada, K7M 5R7. E-mail: christine.doe @ queensu.ca

What has experience got to do with it? An exploration of L1 and L2 Test takers' perceptions of test performance and alignment to classroom literacy activities

Introduction

The prevalence of large-scale testing in K-12 education is well documented in the research literature (e.g., Klinger, DeLuca, & Miller, 2008; Phelps, 2005). Such large-scale tests are typically developed and normed for a first language (L1) test taker population. However, research has demonstrated that second language (L2) test takers may demonstrate differentially lower test performance compared to their L1 counterparts, questioning whether the tests are measuring the same test construct (Abedi, 2002; Fairbairn & Fox, 2009; Fox & Cheng, 2007). One source of this differential test performance may stem from the linguistic complexity in test items, such as word frequency, or passive versus active voice (Abedi, 2002). Beyond linguistic complexity, L2 students may struggle with the genre and register of the test tasks or items. Solano-Flores and Trumbull (2003) called for new research and practice paradigms employing generalizability theory, “to reveal more fine-grained understandings of the interactions among first and second language proficiency, student content knowledge, and the linguistic and content demands of test items” (p. 3).

In the context of the Ontario School Literacy Test (OSSLT), Cheng, Klinger & Zheng (2007) found that two literacy tasks — *narrative* reading and *news report* writing — and two reading skills and strategies — *indirect understanding* and *vocabulary* — were predictors for L1 and L2 test taker group membership. Another factor may be the familiarity that the test takers have with the testing culture (Fox & Cheng, 2007). A recent L2 immigrant student who has no experience taking a test in Canada may be confused about the procedures of the jurisdiction, including test guidelines, time limits, and lack of external support, such as dictionaries or a teacher available for questions. Fox and Cheng concluded that there were three major differences between L1 and L2 test takers: a) knowledge of test taking, b) varied test taker preparation, and c) level of test anxiety. Taken together with the increasing immigration rates of L2 students in Canada (Statistics Canada, 2010), these concerns demonstrate the importance of exploring L2 test takers' perceptions of their test performance and classroom literacy activities in relation to L1 test takers within the context of large-scale testing.

Background

Beyond the types of items used, how test takers approach a test taking situation or the stakes stemming from the test results may influence test performance. There have been a number of studies that have investigated L2 test takers' approaches and experiences with test taking. Such research has focused on the strategies that test takers employ during the test (e.g., Douglas & Hegelheimer, 2007). In addition, test performance is influenced by test anxiety (Horwitz, 2000), background knowledge (Fulcher, 2003), and familiarity with test items (Sasaki, 2000). The relative stakes (high or low) have also been shown to influence the washback of the test (the influence of testing) on students' test preparation and learning (Cheng, 2008, Green, 2006). Despite the number of studies that have investigated test takers' approaches and experiences, there have been few studies to date that have systematically researched both L1 and L2 test taker experiences simultaneously. Identifying the unique features of these two groups of test takers' experience will enhance how the results are interpreted.

Furthermore, the majority of the research exploring test taker experiences has been from one perspective along the test taking process—before or after taking the test. More recently, the testing process is viewed as a dynamic experience where test takers' perceptions evolve (Doe & Fox, 2011; Huhta, Kalaja, & Pitkanen-Huhta, 2006). Doe and Fox observed L2 test takers modify their test taking approach across three low- and high-stakes testing contexts of the same test: within a preparation course, in a practice test, and during the live test. Similarly, Huhta et al. noted variability in students' identities of themselves as test takers within and across various testing events. In our study, we surveyed two groups of test takers' perceptions: one group prior to taking the OSSLT and the other group after taking the OSSLT. We further examined the similarities and differences between L1 and L2 test takers in this testing context.

The OSSLT is designed to measure literacy skills across all the required subject curricula up to Grade 9 (EQAO, 2009a). Despite its cross-curricular intention, L1 and L2 students perceive the OSSLT as an English test (subject-based), or in the case of L2 students, a test of English ability (Fox & Cheng, 2007; Klinger & Luce-Kapler, 2007). Despite the seemingly high stakes of the test as a graduation requirement, the OSSLT is largely a minimum competency test with 84% of the test takers successfully completing the OSSLT (EQAO, 2009a). Although the pass rate for L2 students has increased by 12% since 2006, it is only at 63%, which is 21% below the average (EQAO, 2009a). Based on principals' recommendations, students may defer, rewrite, or satisfy the literacy requirement through alternative means. Deferring the OSSLT is a popular option for students; 32% of L2 students deferred taking the OSSLT in the 2009 administration (EQAO, 2009a). Students who were unsuccessful in passing the OSSLT or have been recommended by their principal may obtain the high school graduation requirement through the successful completion of the Ontario Secondary School Literacy Course (OSSLC) (EQAO, 2009b). These additional options present an equitable approach for accommodating those test takers who otherwise may be at risk of failing the OSSLT. Nevertheless, delays to the successful completion of OSSLT requirement hinder students' course taking options, and future post-secondary plans. Students who complete the OSSLC and are intent on pursuing higher education may be required to take an English language proficiency test for university admittance depending on the number of years of schooling in Canada (see Han & Cheng, in press). Overall, the OSSLT represents a potentially emotionally charged setting for L2 test takers leading to test results that under-represent their true ability, which may not necessarily exist for L1 test takers.

The present study

A survey study was developed based on the literature and the findings from the previous two phases of the research project (Cheng, Klinger, & Zheng, 2007, 2009; Cheng, Fox & Zheng, 2007; Fox & Cheng, 2007; Zheng, Cheng & Klinger, 2007). The first phase explored the differential test performance between L1 and L2 students on the reading and writing sections of the OSSLT. In the second phase, focus groups were conducted with L1 and L2 test takers after completing the OSSLT to compare the two groups' accounts of the test. This study reported here focused on test takers' reflective and retrospective accounts of their upcoming or past test performance and the links between the test and their in-class literacy activities. There were two research questions asked:

- (1) What were L1 and L2 OSSLT test takers' perceptions of their upcoming or past test performance and how did the perceptions vary across the two groups?

- (2) How did OSSLT test takers' perceptions of the alignment between the OSSLT and their classroom reading and writing activities vary by language status and by experience with the test?

Methodology

In order to explore L1 and L2 test takers' perceptions before and after writing the OSSLT, we identified four sub-groups of test takers based on the language status and test-taking experience with the OSSLT: L1 *not taken*, L1 *taken*, L2 *not taken*, and L2 *taken*. The sample for this survey study was obtained from Grades 10 and 11 students in four secondary schools (three public and one private) in eastern Ontario, Canada. These schools were chosen because they had a large cross section of both L1 and L2 students. Of the 528 test takers who participated in the study, there were 213 who had *not taken* the OSSLT (L1 = 141; L2 = 72) and 315 who had *taken* the OSSLT (L1 = 230; L2 = 85). The participants' language status was identified as L2 if the language spoken at home was a language other than English, and if the participants were born in a non-English speaking country. The most commonly reported first languages for L2 test takers were Arabic (7%), Chinese (5%), and French (3%).

Data were obtained from these test takers through a questionnaire composed of three main sections¹. This paper focused on the third section of the questionnaire, with a combination of Likert scale, open-ended questions, and yes/no questions. This section explored test takers' perceptions toward the OSSLT and their test performance, as well as their perceptions of the alignment between the OSSLT and their classroom literacy activities (see Appendix 1).

Data analysis

All of the questionnaire responses were entered into SPSS 16. Test takers were separated through unique identification numbers that allowed us to link the responses to the original questionnaires. Open-coding was applied to the qualitative data to identify meaningful themes within each question over two stages (Patton, 2002). The first stage was to freely generate themes for describing the responses. In the second stage, we looked for "recurring regularities in the data" to reveal the essence of the theme (Patton, 2002, p. 465). We used Excel (2007) to consistently create themes across the responses by question. When there was an inconsistency in coding within or across the questions, either the theme was renamed to be inclusive of the theme in another question or was modified to reflect the themes used within the question. After the first round of coding, an independent researcher recoded approximately thirty percent of the data to provide an additional source of coding reliability. To recode the responses the independent researcher was supplied with descriptions and examples of the themes and an empty coding sheet with test takers' responses. The average reliability co-efficient was 92.6 %. Any inconsistencies were reviewed and modified if the original coding had an error. Once the coding was completed, the themes were linked to the questionnaire identification numbers and merged into the original SPSS file as new variables.

We used quantitative and qualitative approaches to analyze the coded data. Quantitative analyses were first utilized to look for major differences between the language (L1 and L2) and test experience (*taken* and *not taken*) groups. Cross-tabs provided a relative comparison of how these four groups responded by theme. We based any comparisons between the L1 and L2 test takers on percentages due to the uneven sample sizes. For comparison of L1 and L2 test takers

¹ Results from the other two sections of the questionnaire is reported in Zheng et al, (2011).

within the *not taken* or *taken* groups, the percentages were calculated by dividing the count that we were comparing by the L1 or L2 population within that *not taken* or *taken* group. Similarly if we wanted to compare L1 and L2 test takers within a theme, we divided the count by the L1 and L2 population within that theme. In the reporting of the results, we mention the percentage first followed by the count in parentheses. Quotations of individual test takers were used for L1 and L2 test takers to provide a thick description of the themes and presented a rich understanding of the phenomenon (Patton, 2002). Test takers' quotations were chosen if they were characteristic of the L1 and L2 test takers or representative of both groups. Responses are referred by individual identification numbers, appearing in parentheses following a quotation.

To answer the first research question, we compared the themes generated from the *not taken* test takers' perceptions to see if they thought they would pass the test (Q41 and Q52) and the *taken* test takers' reported test results (*pass, did not pass, rewrite*) (Q38 and Q42). The second research question was answered by examining the shifts in the *not taken* and *taken* yes/no responses of whether the test taker saw the test as similar or different to their classroom reading and writing activities (*not taken* Q45 and Q46; *taken* Q47 and Q48). To track the potential shifts in test takers' perceptions, a matrix was created to look at the percentages of coded responses with grouping variables of language status (L1 and L2), test experience (*not taken* and *taken*), test content (reading and writing), and alignment (whether they saw the test as similar or different to classroom literacy activities). Any difference between the *not taken* and *taken* coded responses larger than 5% was recorded. Our decision was not to generalize the findings, but to highlight the differences.

Results

Our results are organized by the two research questions. We began with test takers' perceptions before writing the OSSLT (*not taken* test takers), with a focus on their expectations for the OSSLT, compared with test takers' perceptions after completing the OSSLT (*taken* test takers), with a focus on their test performance. This was then followed by the results on the alignment between the OSSLT and test takers' classroom reading and writing activities.

Test Takers' Perceptions of Test Performance Before and After the Test

A sample of 178 responses were obtained from the two questionnaire items (Q41 and Q52) asked to the *not taken* test takers, and 292 responses were gathered from two questionnaire items (Q38 and Q42) asked to the *taken* test takers. There were four commonly expressed perceptions of future test performance found across the *not taken* test takers: nervous, confident, unconcerned, and unfairness. There were five commonly expressed perceptions of test performance recorded across the *taken* samples: well, not well, bad, unfairness, and indifference. Based on the cross-tab analysis there was a greater difference between L1 and L2 test taker responses in the *not taken* group to the *taken* group. We observed the greatest divergence for *not taken* test takers who expressed being nervous or saw the test as unfair. It should be noted that some test takers expressed nervousness along with confidence, for example, "[I am] a little nervous, but alright overall" (40², L1 test taker). In these cases of mixed nervousness and confidence, we coded the responses as confidence to create the nervousness theme as representative of the test takers who were truly anxious before taking the test. We explored these

² 40 is an individual identification number of a test taker.

responses in more detail to provide further insight into L2 and L1 test taker perceptions before taking the test. In the *taken* group, the ‘well’ and ‘not well’ responses provided contrasting L1 and L2 test taker perceptions after taking the test. Our discussion about the differences and similarities with the test takers responses is coupled with perceptions about whether they thought they would pass or if they had passed the test, respectively, for the *not taken* and *taken* groups.

Overall, more L2 *not taken* test takers expressed being nervous than the L1 test takers [57.6 % (34) L2 test takers compared to 35.3% (42) L1 test takers]. The L1 *not taken* test takers were most likely to express a perception of confidence [35.3% (42) L1 test takers compared to 28.8% (17) L2 test takers]. Interestingly, a greater number of the L1 *not taken* test takers considered the test ‘unfair’ as compared to L2 *not taken* test takers [10.1% (12) L1 test takers compared to 3.4% (2) L2 test takers]. In contrast, the L1 and L2 test taker responses were more similar in the *taken* group. A similar proportion of L1 and L2 *taken* test takers perceived their test performance as ‘not well’ or considered the test as unfair [19% L1 (41) and L2 (14) test takers perceived the test as ‘not well’; 4% (9) L1 test takers and 3% (2) L2 test takers viewed the test as ‘unfair’]. There were however differences observed with those who expressed perceptions that their test performance went ‘bad’ or ‘well’. Proportionately more L2 *taken* test takers viewed their test performance as ‘bad’ [13% (10) L2 test takers compared to 5% (11) L1 test takers]. While more L1 *taken* test takers thought of their test performance as ‘well’ [68 % (147) L1 test takers compared to 60% (45) L2 test takers]. Interestingly, the ‘well’ theme represented the largest category for both groups of *taken* test takers.

There were two key differences recorded for the *not taken* test takers who viewed the test with nervousness. First, a higher proportion of L1 test takers noted time constraints as a concern [12% (4) L1 *not taken* test takers compared to 3% (1) L2 *not taken* test takers]. For example, one L1 test taker commented, “I am nervous. I really want to pass, but I am scared that I won’t. I’m scared that I’ll run out of time (79).” L1 test takers are given 2.5 hours complete the OSSLT and L2 test takers are permitted up to double the time (Education Quality and Accountability Office, 2009a). Second, more L2 *not taken* nervous test takers also alluded to their concern that the test was a test of English [12% (5) L2 *not taken* test takers compared to 9% (3) L1 *not taken* test takers], who perceived their upcoming OSSLT with a sense of nervousness. Within these 8 responses, from both L2 and L1 test takers, all but one referred to the OSSLT as being a measure of English skills. One L2 test taker commented, “I am kind of nervous because it’s a test that is based on reading and writing English, and I am not a superstar in English (34).” In contrast to this skills-based perspective, one L1 test taker related her nervousness to the fact that she saw the OSSLT as an English subject-based test: “I am a bit nervous...[I] don’t have English this semester, [so I] don’t really know if I’ll be prepared enough for it (52).” It is important to note that both L1 and L2 test takers viewed the ‘literacy’ test as a test of English (skill or subject).

We also observed a number of similarities between the L2 and L1 *not taken* test takers, who expressed being nervous about taking the OSSLT. There was a similar proportion of L2 test takers and L1 test takers commented on the high stakes of the test [12% (5) L2 compared to 12% (4) of L1 *not taken* nervous test takers]. One L2 test taker expressed sentiments of being nervous coupled with his concern that he “wouldn’t like to fail the test and then have to take a whole English [OSSLC] course (95).” Here the test taker perceived the OSSLC course as a negative outcome although it was implemented to benefit students, who are struggling with literacy, as identified by failing the OSSLT.

The ‘unfair’ theme presented a blend of nervousness and confidence from the *not taken* test takers (both L1 and L2) in expressing dissatisfaction with the test. Test takers demonstrated

confidence through their strong statements of viewing the test as futile, for example, one L1 test taker stated “I think it’s stupid. It’s a second exam for English that provides no grade.” Not only was the test taker annoyed by having to take a test, we also see this idea again that the OSSLT was considered a test of English. Unique to this group, who viewed the test as ‘unfair,’ were the perceptions that the test was biased or, as one frustrated L1 *not taken* test taker noted, “I heard the pass mark is 60%, so I don’t think it is fair (361).” Test takers seemed angry about the lack of feedback or grades.

Within the unfairness group, there were also fewer test takers who thought they would pass, with 79% predicting they would pass the OSSLT. Overall, this percentage was the lowest for the *not taken* group: 90% of the total *not taken* test taker group, 86% of the nervous test takers, and 100% of the confident test takers thought that they would pass the test. Not surprisingly, the test takers who thought they would fail expressed discontent towards the test, or as one L1 dissatisfied test taker expressed his contempt for the OSSLT and its length, in time and content: “Well, I hate tests, and this test makes students nervous and scared. It is too long and too much to write... (457).” The test takers who viewed the test as unfair represented the most conflicted *not taken* test takers with perceptions that demonstrated nervousness and confidence.

For the *taken* group of test takers, the responses did not vary as much within the themes as was seen in the *not taken* nervous theme. However, when the *taken* responses were paired with the test takers’ reported results of whether they passed the test or not, we were able to gain a deeper understanding of L2 and L1 test takers’ perceptions of their test performance. The highest proportion of test takers who reported passing the OSSLT was the group who perceived their test performance as good compared to the other groups: percentage of *taken* test takers who reported passing the OSSLT were 99% (good), 91% (not well), 90% (indifferent), 73% (unfairness), and 14% (bad). In addition, of the five themes observed for the *taken* group, the test takers who viewed their test performance as bad and with unfairness had the longest responses of the group overall. The longer responses were most likely an indication of the *not taken* test takers’ reflection of their emotional reaction to receiving failing results.

The *taken* test takers who perceived their test performance as unfair expressed a similar dissatisfaction towards the test as the *not taken* test takers. A key frustration for L1 *taken* test takers was not being able to “review [their] mistakes” (123) or feeling that “[the OSSLT] did not reflect [their] true ability in English.” Both L2 and L1 test takers viewed the test as “a waste of time” (118) or “pointless” (353). Once again, issues of feedback and bias came across in the test taker responses as previously seen in the *not taken* group. To further our understanding of L1 and L2 test taker perceptions before and after taking the OSSLT, we probed the perceived connection between the test and the classroom literacy (reading and writing) activities.

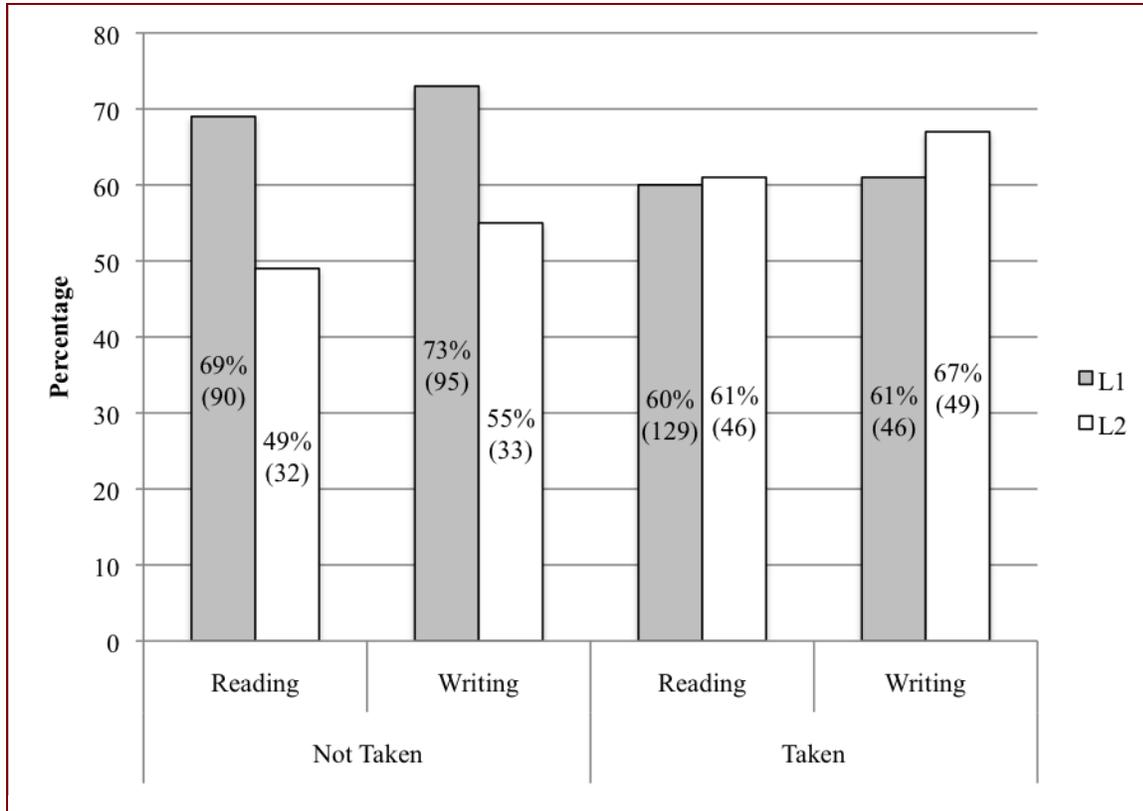
Test Takers’ Perceptions of Alignment Between OSSLT And Classroom Reading and Writing Activities

The *not taken* and *taken* test takers were also asked about their perceptions of whether the OSSLT reading and writing activities were similar to reading and writing activities they completed in class. These perceptions were gathered from 195 *not taken* and 280 *taken* test takers in response to a two-part question (yes/no and open-ended) (*not taken* Q45 and Q46; *taken* Q47 and Q48). From the yes/no portion of the question, more L2 *not taken* test takers compared to L1 *not taken* test takers viewed the reading and writing activities on the OSSLT as different from their classroom reading and writing activities. For the *taken* group, the two language groups were quite similar in their responses (Figure 1). Numerically, the shifts in perceptions between

the *not taken* and *taken* test takers based on language group were approximately 10-13% for both the reading and writing. We then focused on this shift between the *not taken* and *taken* responses and looked to the open-ended responses for insight into this apparent distinction between these two experience groups, especially for the L2 test takers.

Figure 1.

Percentage of L1 and L2 test takers who viewed the reading and writing OSSLT activities as the same to their classroom literacy activities



We noted 11 themes for reading and 12 for writing from the open coding of *not taken* and *taken* responses. The themes related to the test taking process (e.g., reading ability is constant, more strategic), content (e.g., content or format are the same/different), and context of the test (e.g., time and test situation are different) (Tables 1 and 2). Tables 1 and 2 contain the themes for the *not taken* responses and only report the themes for *taken* test takers where there was a change of 5% or more observed.

Table 1.

L1 and L2 test takers' perception of alignment of the OSSLT to their reading classroom activities

	Saw the test as the same			Saw the test as different		
	Coding	L1	L2	Coding	L1	L2
<i>Not Taken Test Takers</i>	No response	33% (44)*	23% (15)	No response	4% (5)	17% (11)
	Always strategic	<.1% (1)	2% (1)	Be more strategic	15% (19)	6% (4)
	Content/Format/ Skills are the same	5% (6)	3% (2)	Content/Format/ Skills are different	4% (5)	6% (4)
	Reading ability is constant	15% (19)	8% (5)**	More difficult	2% (2)	5% (3)
	Preparation	5% (6)	9% (6)	Time/test situation	5.3% (7)	14% (9)
	The same	9% (12)	5% (3)			
	Saw the test as the same			Saw the test as different		
	Coding	L1	L2	Coding	L1	L2
<i>Taken Test Takers</i>	No response	20% (43)	33% (25)	No response	5% (11)	9% (7)
	Always strategic	6%(12)	9% (7)	Be more strategic	7% (15)	5% (4)
	Preparation	3% (7)	1% (1)	Content/format/ skills were different	10% (21)	8% (6)
	Reading ability is constant	14% (30)	3% (2)	Different	5% (10)	1% (1)
				Test situation	7% (15)	9% (7)

*L1 Not taken = 130; L2 Not taken = 65; L1 Taken = 215; L2 Taken = 75.

** 2 L2 test takers reported that the similarity was related to English ability

Table 2.

L1 and L2 test takers' perception of alignment of the OSSLT to their writing classroom activities

		Saw the test as the same		Saw the test as different			
		Coding	L1	L2	Coding	L1	L2
Not Taken Test Takers	No response/N/A		39% (51)*	33%(20)	No response/N/A	6% (8)	20%(12)
	Always strategic		1% (2)	2% (1)	Be more strategic	7% (9)	2% (1)
	Content/Format/ Skills are the same		3% (4)	5% (3)	Content/Format/ Skills are different	2% (2)	5% (3)
	Writing is constant		13% (17)	7% (4)	More difficult	3% (3)	3% (2)
	Preparation		5% (6)	7% (4)	Concrete answers	2% (3)	2% (1)
	The same		8% (11)	7% (4)	Time/test situation	6% (8)	10% (6)
		Saw the test as the same		Saw the test as different			
Taken Test	No response/N/A		26% (55)	36% (26)	No response/N/A	7% (15)	8% (6)
	Content/Format/ Skills are the same		9% (19)	12% (9)	Content/format/ Skills are different	10% (22)	4% (3)
	Preparation		1% (2)	3% (2)	It was easy	3% (6)	1% (1)

*L1 Not taken = 130; L2 Not taken = 60; L1 Taken = 213; L2 Taken = 73.

We observed a pattern in the coded responses similar to the shifts seen in the ‘yes/no’ responses, particularly for the test takers who did not provide a response to the open-ended questions. In reading, there was a matched decrease and respective increase in non-responses for L2 *not taken* test takers who saw the test as different to the *taken* test takers who saw the test as same (approximately 10%). For L1 test takers, we only observed a decrease in reading non-responses for those who saw the test as similar to the yes/no responses (approximately 13%). In writing, we observed a decrease in non-responders for L1 and L2 *not taken* test takers (approximately 10%). Matching increases were not found with the *taken* test taker non-responses for writing.

In reading, we noted a decrease in the percentage of test takers who reported test preparation and a similar increase in the percentage of test takers who reported strategy use as explanations for their perceived alignment. The change was the greatest for L2 test takers: a decrease in preparation (approximately 8%) and an increase in strategy use (approximately 7%). The preparation theme related to the test preparation activities or the time spent studying for the OSSLT test. For instance, one L2 test taker stated, “I did sample tests” (473) and one L1 test taker said, “It’s the same as [the] preparation exercises” (81). This always strategic theme was

characterized as focusing on details or following instructions in both the testing and classroom context with statements like “I read everything carefully” (120, L1 test taker) or “I read the instructions ... [and] underlined an important item from top to bottom” (355, L2 test taker). The potential switch between *not taken* test takers who reported preparation as an explanation to the *taken* test takers who drew on their strategy use as a rationale provides an example of how test taker perceptions may have varied within the two test experience groups.

The only theme that noticeably separated L1 and L2 test takers, regardless of test experience, was the perception that reading and writing ability was constant across the test experience groups. The two groups responded similarly within the theme by indicating that they felt that writing or “reading [was] the same no matter what the circumstance” (L1 test taker, 16). The L1 test takers all mentioned the word reading or writing, while two of the L2 test takers did reason that the similarity was as a result of it being “the same language” (474). Majority of the responses referred to reading as a process or an ability. However, the L1 and L2 test takers commented on how writing styles did not vary by context. For instance, a L1 test taker reported that “I have no reason to change my writing style for the OSSLT” (446). Similarly, a L2 test taker indicated “I can’t change my writing style” (24). Overall, the two language groups felt that reading and writing was considered a process or that writing styles do not vary from context to context. The above results demonstrated that there were many variables which potentially influenced test takers’ perceptions.

Discussion

Our study examined the qualitative responses of a large-scale questionnaire to expand on previous research findings about L1 and L2 test takers’ perceptions of the OSSLT and their test performance and the alignment between the test and their classroom literacy activities, juxtaposed with whether the test takers were preparing to take or had already completed the test.

Test Performance Before and After the Test

We observed a range of perceptions for the OSSLT test takers, with the *taken* L1 and L2 test takers’ perceptions seemingly more similar to each other than in the *not taken* group (RQ1). The *not taken* and *taken* test takers were asked about their perceptions of the OSSLT, specifically, about their respective feelings about taking the test that year or of their past test performance. Since the *not taken* and *taken* test takers were asked slightly different questions (see Appendix 1), the types of responses gathered varied accordingly. The *not taken* responses reflected the test takers’ perceived ability to successfully pass the OSSLT (e.g., nervousness, confidence). The *taken* responses were a judgement of their past test performance (e.g., well, bad). The advantage of asking these two groups slightly different questions was that it allowed us to gather a broader range of perceptions about the OSSLT experience. However, we were unable to look for any generalizations between the *not taken* and *taken* perceptions gathered. Despite this limitation we did observe some possible differences within and across the four test taking groups (L1 and L2; *not taken* and *taken*).

For the *not taken* perceptions, we noted the most variability with test takers who were nervous before taking the OSSLT. The perceptions that were coded as nervousness represented test takers’ anxiety before taking the OSSLT. Similar to previous research, our findings suggest that OSSLT test takers (L1 and L2) are worried about taking the test because of the time constraints, the high stakes of the test, or because they perceive it as an English test—subject or

skill (Fox & Cheng, 2007; Klinger & Luce-Kapler, 2007). We were surprised to observe a greater proportion of L1 test takers express concern about the limited time given to complete the OSSLT. This may reflect the fact that the L2 test takers, who received a time accommodation, felt the time was sufficient or it was not their primary concern. Had we interviewed the L2 *not taken* test takers we would have been able to explore the extent that time constraints influenced their perceptions of nervousness. Regarding the stakes of the test, we observed the same proportion of L1 and L2 test takers perceiving the high stakes as a reason for their nervousness. Beyond simple demographic information, we do not have a lot of information about these L1 test takers and are unsure whether these students were struggling in other aspects of their schooling. The similarity in L1 and L2 test takers' responses does suggest that the OSSLT can be a high-stakes testing context for both language groups, which is likely connected to their nervousness before taking the test. Our findings support the notion that test takers perceive the OSSLT as a test of English (Fox & Cheng, 2007; Klinger & Luce-Kapler, 2007). This perception may reflect how the OSSLT is introduced in schools, which is most likely in English class even though the test designed to be cross-curricular. Further investigation is needed to better understand when and how the OSSLT is introduced to students in their schools.

The fact that proportionally more L2 *not taken* test takers approached the OSSLT with a sense of nervousness and proportionally more L1 test takers expressed a sense of confidence was not surprising. L2 test takers are writing a test focused on reading and writing in their second language while L1 test takers are taking the same test but in their first language. However, with the high percentage of L2 *taken* test takers who reported that their OSSLT test experience went well or that they passed the test, a number of *not taken* L2 test takers were perhaps unnecessarily nervous before the OSSLT. Nevertheless, L2 test takers' uncertainty about their English skills and perceptions of the OSSLT as a test of English would undoubtedly lead to the test anxiety observed.

Perceptions of unfairness were held in common by both *not taken* and *taken* test takers, whether L1 or L2. These perceptions exhibited confidence and nervousness through bold and long answers. This confidence may be masquerading as deep-seated anxiety, especially for the test takers who reported that they thought they would fail or had failed the test. In the *not taken* group, there was a higher percentage of L1 test takers who expressed such unfairness perceptions while this uneven response rate all but disappeared for the *taken* group. There are a couple of explanations for this increasingly similar response rate for the L1 and L2 *taken* test takers. As discussed previously, the two types of questions asked to the test takers elicited slightly different responses. Had we asked corresponding questions we may have observed less separation between the *not taken* and *taken* L2 perceptions of unfairness. However, when we take into account the increased similarity in L1 and L2 test taker responses, another plausible explanation is that L2 *not taken* and *taken* test takers shifted their perceptions of nervousness to unfairness, respectively. Before taking the OSSLT L2 test takers may have been nervous and unsure about their ability to pass, but after experiencing the test they may have felt frustrated about having spent considerable time and energy preparing to take the test and then receiving no feedback. The perceptions gathered from OSSLT test takers about their upcoming or past test performance provided an account of the test takers experiences. To complement these highly personal perceptions, we looked to the test takers' more evaluative perceptions of the alignment between the test and their classroom literacy activities.

Alignment between OSSLT and Classroom Reading and Writing Activities

The exploration of the grouping variables—language status, test experience, test content (reading and writing), and alignment (whether test takers saw the test as similar or different to classroom literacy activities)—together answered the second research question. If one of the factors was removed from the analysis, we would have had a distorted view of test takers' perceived alignment between the OSSLT and their classroom literacy activities. Surprisingly, test taker perceptions varied by whether they provided an explanation to their initial 'yes/no' response. The fact that we only observed an increase for reading and not for writing L2 *taken* non-responses suggests that the writing portion of the test was more memorable for the L2 test takers. A possible explanation is that in the writing section test takers must produce something that takes time to construct and requires more creativity than responding to questions about a reading passage.

A potential shift from test preparation to strategy use suggests that taking the test provided L2 test takers with the opportunity to recollect what they did on the test and the actual strategies they employed versus how they prepared for the test with their teacher and guessed the contents of the test to be the same or different. The most noticeable separation between L1 and L2 test takers' perceptions was reading and writing were seen as constant abilities. The perception of reading or writing as a constant skill demonstrates the L1 test takers' confidence about their ability, regardless of experience, while it seems that L2 test taker perceptions about their ability were influenced by the test experience.

Conclusion

In addition to the limitations already mentioned, we could not track test taker responses along the testing process. Our data forming the *not taken* and *taken* perceptions were gathered from two independent groups of test takers at one time. As such, the findings should be interpreted with caution. Nevertheless, the findings highlight the need for more research on the effect of test experience on test taker perceptions. In addition, we were unable to link the questionnaire responses to the testing outcomes. Majority of the *taken* test takers reported passing, but we do not know if the L2 test scores were markedly lower than L1 test takers as was seen in previous studies on the OSSLT or other tests developed for an L1 test population (Abedi, 2002; Cheng et al., 2007, 2009; Fairbairn & Fox, 2009). Had we been able to connect the actual test scores to the perceptions gathered we would have been able to uncover more of the complexities inherent to OSSLT test taker perceptions. Indeed, our data source was from the test takers themselves through self-reported questionnaire responses. However, such a large-scale tool allowed us to capture the perceptions from both *not taken* and *taken* as well as L1 and L2 test takers. This matrix of perspectives allowed us to gain a broader insight into the OSSLT test experience.

We were able to obtain a sampling of four schools that were representative of the total OSSLT test population. Perhaps not represented in our sample were the L2 students whose parents are not literate in English and did not sign the required active parental consent for their child to participate in our study. More research on literacy in schools is needed that considers the students and their parents who have limited literacy experience at home, whether it be in their first language or English. Typical recruitment strategies, especially those employed in large-scale studies, often do not connect with this neglected population because of their limited literacy skills.

Despite these limitations, the findings highlight the necessity to factor in multiple perspectives, not only L1 and L2, but also from test takers at different points along the testing process. Most salient was the role that test experience had on test takers' perceptions of their test performance and the alignment between the test and their classroom literacy activities. Taking the test was influential for both L1 and L2 test takers, albeit, greater for the L2 group. This effect of test experience suggests that test takers would benefit from improved test preparation before taking a large-scale test like the OSSLT. Such test preparation activities should provide opportunity for students (both L1 and L2) to realize that they have the ability to pass the test, thereby alleviating any nervousness they may have before taking the test or highlighting where they are weak and need to focus their studies.

References

- Abedi, J. (2002). Standardized achievement tests and English language learners: Psychometric issues. *Educational Assessment*, 8(3), 231-257.
- Cheng, L. (2008). Washback, impact and consequences. In E. Shohamy and N. H. Hornberger (Eds.), *Encyclopedia of Language and Education, Vol. 7: Language Testing and Assessment* (2nd ed., pp. 349-364). New York: Springer Science+Business Media.
- Cheng, L., Klinger, D., & Zheng, Y. (2007). The challenges of the Ontario Secondary School Literacy Test for second language students. *Language Testing*, 24(2), 185-208.
- Cheng, L., Klinger, D., Zheng, Y. (2009). Examining students after-school literacy activities and their literacy performance on the Ontario Secondary School Literacy Test. *Canadian Journal of Education*, 32(1), 118-148.
- Cheng, L., Fox, J. & Zheng, Y. (2007). Student accounts of the Ontario Secondary School Literacy Test: a case for validation. *Canadian Modern Language Review*, 64(1), 69-98.
- Doe, C., & Fox, J. (2011). *Exploring the testing process: Three test takers' observed and reported strategy use over time and testing contexts*. *Canadian Modern Language Review*, 67(1), 29-54.
- Education Quality and Accountability Office (EQAO). (2009a). *Guide for accommodations, special provisions, deferrals, and exemptions: Support for students with special education needs and English language learners*. Toronto, ON: Queen's Printer for Ontario. Retrieved from <http://www.eqao.com/Educators/Secondary/10/10.aspx?Lang=E&gr=10>
- Education Quality and Accountability Office (EQAO). (2009b). *Ontario student achievement: EQAO's provincial report on the results of the 2008-2009 Ontario School Literacy Test*. Retrieved from http://www.eqao.com/pdf_e/09/CPRR_Xe_0609_WEB.pdf&sa=X&ei=wiMxTNLdC9GFnQfN66ndAw&ved=0CBgQzgQoADAA&usg=AFQjCNH0KXgDghlwXraJKZmtDu7PtBhJlg
- Douglas, D., & Hegelheimer, V. (2007). Assessing language using computer technology. *Annual Review of Applied Linguistics*, 27, 115-132.
- Fairbairn, S. B., & Fox, J. (2009). Inclusive achievement testing for linguistically and culturally diverse test takers: Essential considerations for test developers and decision makers. *Educational Measurement: Issues and Practice*, 28(1), 10-24.
- Fox, J. & Cheng, L. (2007). Did we take the same test? Differing accounts of the Ontario Secondary School Literacy Test by first and second language test takers. *Assessment in Education: Principles, Policy & Practice*, 14(1), 9-26.
- Fulcher, G. (2003). Task difficulty in speaking tests. *Language testing*, 20(3), 321-344.
- Green, A. (2006). Washback to the learner: Learner and teacher perspectives on IELTS preparation course expectations and outcomes. *Assessing Writing*, 11, 113-134.

- Han, H. & Cheng, L. (in press). Tracking the success of ESL students within the context of the Ontario Secondary School Literacy Test. *Canadian and International Education Journal*, 40(1), 76–96.
- Horwitz, E. K. (2000). Horwitz comments: It ain't over 'til it's over: on foreign language anxiety, first language deficits, and the confounding of variables. *The Modern Language Journal*, 84(2), 256-259.
- Huhta, A., Kalaja, P., & Pitkänen-Huhta, A. (2006). Discursive construction of a high-stakes test: The many faces of a test-taker. *Language Testing* 23, 326-350.
- Klinger, D. A., & Luce-Kapler, R. (2007). Walking in their shoes: Students' perceptions of large-scale, high-stakes testing. *Canadian Journal of Program Evaluation*, 22, 29-52.
- Klinger, D., DeLuca, C., & Miller, T. (2008). The evolving culture of large-scale assessments in Canadian education. *Canadian Journal of Educational Administration and Policy*, 76, 1-34.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Phelps, R. P. (2005). *Defending standardized testing*. Mahwah, NJ: Lawrence Erlbaum.
- Sasaki, M. (2000). Effects of cultural schemata on students' test-taking processes for cloze test items: a multiple data source approach. *Language Testing*, 17(1), 85-114.
- Solano-Flores, G., & E. Trumbull. (2003). Examining language in context: the need for new research and practice paradigms in the test of English-language learners. *Educational Researcher* 32, 3–13.
- Statistics Canada. (2010). *Canada's population estimates*. Retrieved from <http://www.statcan.gc.ca/daily-quotidien/100628/dq100628a-eng.htm>
- Zheng, Y., Cheng, L. & Klinger, D. (2007). Do test formats in reading comprehension affect ESL/ELD and non-ESL/ELD students' test performance differently? *TESL Canada Journal*, 25(1), 65-80.
- Zheng, Y., Klinger, D., Cheng, L., Fox, J., and Doe, C. (2011). Test-takers' background, literacy activities, and their views of the Ontario Secondary School Literacy Test. *Alberta Journal of Educational Research*, 57(2), 115–136.

Appendix 1: Questionnaire items reported in this study***Both Taken and Not Taken Questionnaires***

Q10a: What language did you **first** speak at home?

Q10b: Where were you born?

Not Taken Questionnaire

Q41: Do you expect that you will pass the OSSLT?

Q45: Do you think the way you read on the OSSLT will be the same as the way you read in your classes? Yes/No. Please explain.

Q46: Do you think the way you write on the OSSLT will be the same as the way you write in your classes? Yes/No. Please explain.

Q52: How do you describe your feelings about taking the OSSLT this year (e.g. nervous, unconcerned)?

Taken Questionnaire

Q 38: Which of the following best describes your results on the OSSLT? I passed the test; I did not pass the test and will rewrite it; I did not pass the test and will take the OSSLC instead.

Q42: How do you feel about your test performance?

Q47: Do you think the way you read on the OSSLT was the same as the way you read in your classes? Yes/No. Please explain.

Q48: Do you think the way you wrote on the OSSLT was the same as the way you write in your classes? Yes/No. Please explain