

---

# Teachers' Assessment of ESL Students in Mainstream Classes: Challenges, Strategies, and Decision-Making

*Terry Milnes and Liying Cheng*

---

*Given the increasing numbers of ESL students in Canadian classrooms, this study investigated how teachers of mainstream classes assess the written work of ESL students and whether they use different assessment strategies for ESL versus non-ESL students. Interviews were conducted with seven mainstream teachers from a private high school in Ontario. Although within-school variation was evident in the participants' approaches to assessing the work of both ESL and non-ESL students, most participants modified their assessment strategies when marking the work of ESL students. This finding suggests a need for school-level discussions and structured professional development activities relating to assessing ESL students' work.*

*Face à l'augmentation du nombre d'élèves en ALS dans les écoles au Canada, nous avons étudié la façon dont les enseignants dans les classes ordinaires évaluent le travail écrit des élèves et dans quelle mesure ils emploient les mêmes stratégies d'évaluation pour eux que pour les élèves qui ne sont pas en ALS (non ALS). Sept enseignants de classes régulières dans une école secondaire privée en Ontario ont participé à des entrevues. Alors qu'une variation individuelle s'est manifestée dans les approches à l'évaluation des élèves en ALS et des élèves non ALS, la plupart des participants ont indiqué qu'ils modifieraient leurs stratégies quand ils évaluaient le travail d'élèves non ALS. Ces résultats donnent à penser qu'il serait nécessaire pour les écoles d'avoir des discussions et de structurer des activités professionnelles portant sur le travail des élèves en ALS.*

---

## Introduction

Census and immigration statistics reveal that in 2002 almost 46% of immigrants to Canada spoke neither English nor French as a first language (Citizenship and Immigration Canada, n.d.). Canadian high schools, particularly in urban centers, have been described as "transformed" (O'Byrne, 2001) by steadily increasing numbers of English as a second language (ESL) students. In Ontario, provincial curriculum and policy documents instruct high school teachers to make "appropriate modifications to teaching, learning, and evaluation strategies ... to help students gain proficiency in English" (Ministry of Education, 2000, p. 10); these documents further note that

“teachers of all subjects are responsible for helping students develop their ability to use English” (p. 10).

There has been little research into how (and indeed whether) Canadian high school teachers modify instruction and assessment practices with respect to the growing cohort of ESL students in mainstream Canadian classrooms. Comparatively more research studies have been conducted in the United Kingdom, Australia, and Hong Kong in the context of English as an additional/second language (Rea-Dickins, 2004; Davison, 2004, 2006; Leung & Mohan, 2004) and in the higher education context (Cheng, Rogers, & Hu, 2004; Cheng, Rogers, & Wang, 2008; Erdösy, 2007). The current study was developed in response to a scarcity of Canadian-based research in the high school setting. It focused on the assessment-related practices of seven Ontario high school teachers whose mainstream classes included ESL students.

### **Teachers’ Assessment Practices and ESL Students in Mainstream Classrooms**

Two major areas of the literature provided the theoretical framework for this study: research relating to teachers’ classroom assessment practices; and studies relating to ESL students in mainstream classrooms.

#### *Teachers’ Assessment Practices*

Research into classroom assessment practices suggests that there is a divergence between approaches recommended by measurement specialists and those actually used by classroom teachers (Stiggins, 2001). Cizek, Fitzgerald, and Rachor (1996) noted variation in United States elementary and high school teachers’ assessment practices, but were unable to identify predictors of this variation such as sex, years of teaching experience, setting, or degree of familiarity with district assessment policies. A “success bias” (p. 175) was built into teachers’ assessment practices: final grades of their students appeared to represent a combination of formal and informal achievement measures as well as nonachievement factors. On the basis of interviews with teachers from Virginia, most of whom taught at the high school level, McMillan and Nash (2000) characterized classroom assessment practices as individualized and idiosyncratic. Teachers’ practices were influenced by an inclination (among others) toward “pulling for students” (p. 9). Further, teachers tended to reward perceived student motivation and effort in their assessment of students’ work. After surveying middle and high school teachers in Virginia, McMillan (2001) identified “academic enablers” (p. 28)—including effort, ability, improvement, and participation—which along with achievement, external benchmarks, extra credit work, and consideration of borderline cases were distinct components of teachers’ grades. Within-school variance in teachers’ grading practices was found to be greater than between-

schools variance in a study involving elementary school teachers in seven Virginia school districts (McMillan, Myran, & Workman, 2002). The same study noted the complexity of teachers' notions of effort, suggesting that improvement and ability may be part of this construct.

Airasian and Jones (1993) characterized teachers' real-world assessment decisions as "messy" (p. 242), involving choices from among a range of reasonable alternatives. The researchers pointed to the tension inherent in the expectation that teachers individualize instruction based on their knowledge of their students and yet assess the same students' work dispassionately. According to the authors, teachers focused on fairness in assessment (including consideration of students' needs and circumstances) rather than objectivity. Brookhart (1994) reviewed a number of earlier studies, finding that teachers often considered effort and ability as well as achievement when assigning grades; that elementary school teachers emphasized observation and informal measures of achievement, in contrast to high school teachers' tendency to rely on written measures of achievement; and that there was variation in teachers' grading practices. Finally, Brookhart noted that "teachers clearly consider grade uses and consequences when they assign grades" (p. 291); thus teachers were concerned about the consequential validity of the marks they awarded (Messick, 1989).

### *ESL Students in Mainstream Classrooms*

Literature relating to the presence of ESL students in mainstream classes has suggested that some teachers view subject-related content mastery and second-language development as separate processes, with the latter seen as the responsibility of ESL-specialist teachers (Penfield, 1987). Ruiz-de-Velasco and Fix (2000) made a similar finding in a study of ESL students attending high schools in impoverished areas of US cities: they reported that teachers of mainstream classes felt intimidated by the task of developing strategies of benefit to ESL students. Mainstream teachers who had had limited contact with ESL students often had low expectations of these students; teachers' attitudes were more positive in schools where with the encouragement of principals, ESL-specialist teachers and teachers of mainstream classes collaborated. In a UK study set in three primary schools, Franson (1999) reported that teachers of mainstream classes expressed confidence in their ability to work with ESL students, but spoke about instructional and assessment issues related to ESL students in general terms: teachers were comfortable with the presence of ESL students in their classes, but did not know about specific instructional and assessment strategies that might benefit these students.

In a Canadian study, Klesmer (1994) found that teachers overestimated the achievement of 12-year-old ESL students on a set of reading, writing, speaking, and listening tasks. He noted that particularly with respect to

listening and speaking tasks, teachers mistook competence in everyday communication skills as language skills that were more highly developed than objective test instruments suggested (see also Cummins, 1996). A consequence of this finding was the possibility of classroom teachers failing to identify at-risk ESL students. Klesmer hypothesized that teachers may have given ESL students the benefit of the doubt during the assessment process; or that teachers did not have fact-based or even intuitive norms for ESL students' achievement based on students' age and length of residence in Canada.

Although Penfield (1987) and Ruiz-de-Velasco and Fix (2000) found a separation between the roles of content-area and ESL-specialist teachers, recent research literature also reflects examples of integrated language and content-area instructional approaches for high school teachers (Schleppegrell & de Oliveira, 2006; Hammond, 2006). With respect to classroom assessment, Huang and Morgan (2003) noted the increasing attention paid by some teachers of mainstream classes to combining language and content in instruction, but pointed to difficulties teachers apparently have in integrating language and content when assessing ESL students' written work. Working in a framework of functional linguistics and knowledge structures, the researchers studied successive drafts of an exercise in classification produced by ESL students in two Canadian high school science classes. Students received support in both the science content and the linguistic aspects of producing writing related to classification. Careful evaluation and comparison of students' first and third drafts suggested the interrelatedness of content instruction and language development. Although Huang and Morgan acknowledged that few high school teachers currently have enough linguistic knowledge to evaluate ESL students' writing in terms of both discipline-specific content and discourse features, they nonetheless pointed to efforts currently underway to develop resource materials that will assist teachers in recognizing the language features associated with their own academic disciplines. Such materials will help high school teachers integrate the assessment of content and language for ESL students.

In another Canadian study, Erdösy (2007) investigated the feedback provided by a Canadian university professor to undergraduate students in a class of both ESL and non-ESL students. Although this research took place in the context of postsecondary rather than high school education, a finding emerged that is relevant to researchers working in both settings. Erdösy noted that feedback to students must be understood and analyzed not only through specific written comments made by assessors on test papers or essays, but also in the larger context of in-class discussions of course assignments and expectations.

This review of the research literature suggests the need for more empirical studies to be conducted, particularly in Canada, to explore the factors that

influence assessment decision-making on the part of mainstream teachers of ESL students. Our study centered on the following research questions: How do high school teachers assess the written work of ESL versus non-ESL students in mainstream classes? What relative weights do teachers of mainstream classes assign to achievement and nonachievement factors when assessing the written work of ESL students? How does this compare with the weighting of achievement and nonachievement factors assigned by the same teachers when they assess the written work of non-ESL students?

## **Methodology**

We investigated these questions through the approach of naturalistic inquiry. We gathered data through the use of in-depth interviews guided by (but not limited to) a list of questions reflecting both the research literature and the research questions underpinning the study. This semistructured interview approach was chosen because it ensured a degree of consistency in the data we collected while also allowing for the use of probes to invite participants to elaborate on their attitudes and experiences (Patton, 2002).

### *Research Site*

The research site was a private school in Ontario to which we have assigned the pseudonym of Ontario Private School (OPS). One third of the high school students were international students from Hong Kong, Korea, Japan, Taiwan, Indonesia, Venezuela, Mexico, and Germany. Although they were not true beginners in their ESL studies, these international students did not speak, read, or write English fluently; nonetheless, they planned to attend Canadian or US colleges and universities on graduation from high school. They enrolled in ESL classes at OPS and often supplemented their studies with private tutoring. ESL students at OPS were mainstreamed relatively quickly: sheltered instruction in core subjects such as science, history, and geography was not offered beyond the grade 10 level. As a result, grades 11 and 12 classes at OPS were composed of both ESL and non-ESL students. In this study, only ESL and FSL (French as a second language) courses were defined as other than mainstream classes. Thus mainstream classes at OPS included history, geography, English, mathematics, science, health and physical education, civics, the arts, computing, and business studies.

### *Participants*

Because this study focused on the assessment practices of teachers of mainstream classes, it was this cohort of teachers (rather than ESL or FSL teachers) that formed the pool of potential participants for our investigation. At the time the study was undertaken, all potential participants had had at least one full year of teaching experience at OPS and as a consequence had made assessment-related decisions for both ESL and non-ESL students that

ranged from grading individual assignments to marking final exams and making end-of-year pass/fail decisions. Teachers of mainstream classes during the academic year 2003-2004, therefore, formed a convenience sample for the purposes of this study, and all were invited to participate. Seven teachers agreed to do so, of whom two were female and five were male. Pseudonyms were assigned to all the participants. Ross, with fewer than five years of teaching experience, taught history and civics. Eric had taught mathematics, careers, and business and information technology over the course of more than five years as a high school teacher. Douglas reflected on his career of more than 35 years as an English teacher; his colleague Anton, also an English teacher, had 20 years of experience as an educator. Faith had taught biology, science, and geography during the course of a teaching career spanning more than 25 years. William was a history and careers teacher with just over five years of teaching experience. The final participant was Maria, who had been a health and physical education teacher for slightly less than five years. Collectively, the participants reported assessing a variety of types of written work by students including lab reports, expository paragraphs and essays, projects, research reports, letters, tests, and examinations.

### *Data Collection and Analysis*

Individual interviews with each of the participants were conducted from August to October 2004; these interviews were approximately one hour in length and were audiotaped (see Appendix for interview questions). The tone of the interviews was relaxed, with participants discussing their assessment-related experiences at some length and often supplementing their responses with anecdotes and references to specific writing tasks or to individual students. All the interviews were conducted by one researcher who generated field notes after each interview.

The interviews were transcribed verbatim, segmented, and labeled according to the data analysis procedures outlined by McMillan and Schumacher (2001). Topic labels were generated from the participants' comments as well as from the literature review; these topics were then coded and collated. Data analysis involved an iterative process of examining the collated data segments, field notes, and interview transcripts; grouping topics into categories and sub-categories; and identifying themes in the participants' comments. These themes were used to build a portrait of the assessment-related practices that the participants reported using at OPS.

### **Findings and Discussion**

Data analysis suggested a number of findings related to the assessment decisions made by these participants, all of whom were experienced in the task of assessing the written work of both ESL and non-ESL students. We consider each of these findings in turn.

### *Participants' Views of ESL Students and ESL Students' Written Work*

A distinction emerged from the data between the participants' views of ESL students and the writing these students produced. ESL students themselves were generally referred to positively by the participants, most often being described as hard-working. Relative to their non-ESL peers, ESL students were also described as more likely to complete homework assignments (Ross); to articulate strong moral values (Eric); to attend teacher-led tutorial sessions (Douglas); and to attempt longer answers to test questions (Maria). Lack of participation in class discussions was one of the few negative comments made about ESL students, but the participants offered explanations for this phenomenon, referring to shyness on the part of ESL students (Faith); lack of confidence in their speaking skills (Eric); and the possibility that some discussion topics made ESL students feel uncomfortable (Maria). However, for the most part the participants related anecdotes of ESL students' successes in the face of the academic, cultural, and linguistic challenges that they faced. Douglas noted that the work ethic that ESL students evinced "raised the bar" for all students in his mainstream classes. William remarked, "One has to take one's hat off to students who are learning in a second language. The amount of effort and intelligence that is required to do that is immense."

Each participant readily outlined characteristics of good writing for the subject area in which he or she taught, however, and the written work submitted by ESL students often (although not always) fell short of the participants' expectations with respect to writing style. For most of the participants, this raised an issue captured by Eric: "Do we mark the work [ESL students] produce? Or do we mark the learning that is taking place?" Eric's response to this issue was to assess the work of ESL students in a wider context of learning and growth, processes in which he viewed most of his ESL students as actively engaged.

For at least two of the participants, there was a degree of stress associated with grading ESL students' writing. Ross used the words *dilemma* and *quandary* when discussing his marking process for ESL students' work:

I know what the [ESL] student is trying to say but they haven't said it which sort of puts me into a bit of a quandary. Do I assess them based on what I think they're trying to say or what they've actually written?

Maria's experience was that the writing skills of both ESL and non-ESL students were generally weaker than she expected. However, she felt that some ESL students

were so disadvantaged that they couldn't often even communicate. Their sentences were so fragmented and mixed up, they didn't even make any sense. And so, [marking] became like an exercise in futility

sometimes, depending on their level of ESL. And I always found that really hard.

Thus the participants evinced respect for the ESL students in their classes, but at the same time identified a number of assessment-related issues relating to this cohort of students.

### *Participants' Use of Differentiated Assessment Approaches for ESL Versus Non-ESL Students*

The within-school variation in teachers' assessment approaches on which McMillan et al. (2002) commented was evident among the participants in this study. Table 1 summarizes the approaches that the participants reported using.

Alone among his colleagues, Douglas commented that he adhered strictly to rubric criteria when assessing students' written work regardless of whether students were ESL or non-ESL. Errors made by ESL students cost them marks on style-related rubric elements, and thus these students generally received lower marks in Douglas' classes than did non-ESL students. The other participants in this study reported adjusting their assessment approaches for ESL students, however. Based on their awareness of the challenges faced by ESL students who were acquiring a new language in addition to subject-specific content, Eric and William lowered their expectations of style when assessing the written work of ESL students. Three of the participants (Ross, Faith, and Maria) alluded to interpreting test and examination answers written by ESL students, searching for evidence of ESL students' understanding of curriculum concepts. Faith also noted that she allowed her ESL students extra time to complete homework assignments. Anton stated that he simply did not apply the same set of expectations to English essays submitted by ESL students as he did to essays from non-ESL students; as a consequence, he reported acknowledging through marks what he characterized as "small steps" forward on the part of ESL students. Maria adapted some assignments for ESL students, allowing them to present information visually rather than in written format.

In spite of adjustments in the assessment strategies adopted by most of the participants, ESL students were nonetheless perceived to remain at a disadvantage relative to their non-ESL peers. Faith noted that ESL students did not generally have the time to complete extra credit work for additional term marks, for example. Eric commented on the impersonal and mechanical nature of ESL students' writing and suggested that at times this may have had a negative effect on the marks he awarded. Maria estimated that the written work submitted by her ESL students lowered their overall course mark by as much as 10 or 20%.

The portrait that emerges from the participants' comments about their assessment practices with respect to ESL students is thus complex. At times

Table 1  
Participants' Assessment Approaches

	<i>Mainstream Classes Taught and Discussed During Interview</i>	<i>Nonachievement Factors Influencing Assessment for ESL and Non-ESL Students</i>	<i>Adjustments in Assessment for ESL Students</i>
<i>Ross</i>	history, civics	homework, class participation, effort	interpreting test answers
<i>Eric</i>	mathematics, careers, business information, and technology	effort (all grades), homework completion (grades 9 and 10)	lowering style expectations relative to non-ESL students
<i>Douglas</i>	English	none	none
<i>Anton</i>	English	effort, class participation	using marks to reward small steps in progress
<i>Faith</i>	biology, science, geography	homework, class participation	substituting effort for class participation; allowing extra time to complete homework assignments; interpreting test answers
<i>William</i>	history, careers	effort (history and careers), class participation (careers only)	rewarding growth and effort through marks; lowering style expectations relative to non-ESL students
<i>Maria</i>	health and physical education	in health education component: effort	in health education component: adapting assignments; interpreting test answers

ESL students received a number of considerations that non-ESL students did not, including lower teacher expectations of writing style than for non-ESL students as well as the interpretation of ESL students' written work. Even given these considerations, however, many of the participants viewed ESL students as being at a disadvantage compared with non-ESL students in the same classrooms.

#### *The Role of Achievement and Nonachievement Factors in Participants' Assessment Decisions*

The assessment research literature suggests that the marks awarded by classroom teachers reflect a mix of achievement and nonachievement measures (Brookhart, 1994; Cizek et al., 1996; McMillan, 2001). This was the case for the participants in this study, again with the exception of Douglas. Report card marks for both ESL and non-ESL students in Ross' history and civics classes, for example, included components for effort, participation in class discus-

sions, and homework completion. Ross' colleagues included the following nonachievement elements for ESL and non-ESL students when grading assignments or arriving at term and final marks: effort (Eric, Anton, William, Maria); participation in class discussions (Anton, Faith, William); and homework completion (Faith and for his grades 9 and 10 students, Eric).

Because the participants generally used a case-by-case approach to assessment decisions (see below), they did not quantify the effect of nonachievement factors on report card marks. For all the participants except Douglas, however, the effect of nonachievement factors on students' marks was positive: generally speaking, students' marks were, in the words of Anton, "bumped up" by the effect of nonachievement factors. This is consistent with the "success bias" referred to by Cizek et al. (1996, p. 175) as well as with McMillan and Nash's (2000) finding that classroom teachers are inclined toward "pulling for students" (p. 9).

The Ontario Provincial Report Card (Ministry of Education, Ministry of Training, Colleges and Universities, 2005) separates the reporting of achievement from nonachievement factors, but this distinction was disregarded by most of the participants in the current study. This may have reflected a belief on the part of the participants that learning encompasses the acquisition of work habits and attitudes as well as the understanding of curriculum concepts. It may also have related to Brookhart's (1994) note that the consequential validity of the marks they award is important to teachers. At OPS, for example, marks alone determined students' placement on the school Honour Roll; further, in the Ontario setting, report card information about marks, and not about students' learning skills, is submitted to the province's centralized and computerized university application center (Ontario Universities' Application Centre, 2005).

#### *The Role of Student Effort in Participants' Assessment Decisions*

With respect to the role of nonachievement factors in assessment, one additional theme that emerged from the participants' comments was that they generally valued and rewarded the nonachievement factor of effort more than any other. This occurred with work submitted by both ESL and non-ESL students. Eric noted:

I think what I have been guilty of is I will mark a paper and it could be argued it would be mediocre or poor quality. But because of my personal relationship with the student, because of anecdotal evidence, because of past observations, because of time over coffee in the dining room working on this exercise, I just have a really comfortable grasp that they have put their best effort into this. I find myself wanting to reward that. I find myself wanting to raise the grade.

There was variation in the extent to which perceived student effort influenced marks, however. For example, some participants (Eric, Faith) reported that effort might influence a pass/fail decision at the end of the year if a student's mark was just below the pass threshold of 50%. Even Douglas, who otherwise provided data that differed from information provided by his colleagues, commented that he perceived and valued evidence of student effort although he did not reward this nonachievement factor directly through the marks he awarded.

McMillan et al. (2002) alluded to the complex nature of effort, suggesting that it is a construct that might be mediated by teachers' perceptions of ability and improvement. The potentially multifaceted nature of the construct of effort was underscored in this study by the variety of indicators of student effort referred to by the participants. It appeared that they did not simply equate *effort* with a vague notion of *working hard*, but rather gauged student effort on the basis of specific indicators. These indicators included: writing longer essays or answers to test and examination questions (Eric, Douglas, Maria); attending teacher-led tutorial sessions (Douglas, Faith); using dictionaries (Anton); working conscientiously during class (Ross); writing multiple drafts of essays and revising each one carefully (Anton); conferring with teachers about projects and assignments (Eric, Maria); jotting notes in student-owned textbooks (Anton); and handing in assignments and projects that were neat in appearance (Maria). Of note is that several of the participants mentioned indicators of effort specific to ESL students, including speaking English in the school hallways between classes (Anton); working with private tutors (Eric); appearing tired (Faith); carrying and consulting a thesaurus (Anton); and using varied sentence structure, even at the risk of making errors, when revising essays (Anton). These findings suggest that research into the role of effort in teachers' assessment decisions will need to acknowledge both the complex nature of the construct of effort and the possibility that teachers may apply differentiated measures of effort to different groups of students.

### *Resolving Assessment Decisions*

If many of the participants reported facing (with varying degrees of stress) complex decisions when assessing ESL students' written work, how did they go about resolving those decisions? Apart from Faith, who routinely substituted effort for class participation marks for ESL students, and Douglas, who adhered firmly to his rubric, the participants did not report applying specific or consistent policies when assessing ESL students' writing. Indeed, it appeared that most of the participants adopted a case-by-case approach to assessing the writing of both ESL and non-ESL students. This is consistent with Airasian and Jones' (1993) suggestion that teachers find it difficult to set aside their knowledge of individual students and become impartial judges of

achievement. Eric, Anton, Faith, and William, in fact, explicitly rejected the possibility of teacher objectivity in assessment decision-making. Their knowledge of individual students, gleaned through multiple opportunities to observe and interact with learners in the small classes and shared lunches characteristic of this private school setting, may also have contributed to the participants' comfort level in adopting this case-by-case approach to assessment. In many respects, the ways that many of the participants at OPS combined informal observations with formal measures of achievement to make assessment decisions were similar to the practices adopted by elementary school teachers (Brookhart, 1994).

Although the assessment practices reported by the participants may indeed have been idiosyncratic and individual (McMillan & Nash, 2000), these practices were not arbitrary. The participants in the current study appeared to be actively engaged in seeking comprehensive evidence of student learning. For example, both Eric and Faith reported supplementing ESL students' written work with conversations aimed at ascertaining whether students had indeed grasped key curriculum concepts, but were struggling to express their understanding in written form. The participants were guided in their assessment decisions about ESL students' work by their own beliefs and concerns about such issues as how best to motivate students and how to be fair to all students. ESL students' grade level and the participants' notions of academic standards also influenced the resolution of assessment decisions, with most participants reporting that their expectations for the work of all students (both ESL and non-ESL) increased in grade 11 and particularly in grade 12. Standards, coupled with the stated mandate of OPS to prepare its graduates for postsecondary education, trumped the sympathy some of the participants felt for ESL students who grappled with curriculum content in a second language. Ross, for example, commented,

And I have had that challenge where a student that I had just did not have the English level to understand the course material but worked hard at it but still ended up with a failing result in the course.... sometimes they would be answering something other than what I asked, and that really spoke loudly to me and clearly to me that they were not understanding the English at a level where they should have been in that course.

Another factor relevant to resolving assessment decisions about ESL students' writing was the subject taught by the participants. Eric and William both noted, for example, that had they been English teachers they would have had higher expectations for the written work submitted by their ESL students. Thus it appeared that most of the participants resolved assessment dilemmas through a complex process in which they balanced their knowledge of individual students with beliefs about such issues as academic

standards and the relative importance of writing skills in the disciplines in which they taught.

### *Participants' Uncertainty about Their Role in Assessing and Promoting Language Growth*

Although several of the participants made it clear that they valued evidence of growth in student learning, Eric commented on two types of growth that he observed occurring in his ESL students: first, growth in mastery of subject content; and second, growth in language development. He felt that he could influence only one type, namely, growth in understanding of mathematics concepts. Growth in language acquisition was described by Eric as something ESL students "seem to go away and do on their own." Both Maria and Ross alluded to the language acquisition needs of their ESL students, but did not outline or express awareness of specific strategies to meet these needs. This is perhaps not surprising given that the participants had not had extensive inservice sessions relating to stages of second-language development, the identification of at-risk ESL learners (Klesmer, 1994), or research-based instructional and assessment approaches of particular benefit to ESL students. Instruction in language learning was apparently seen largely as the bailiwick of ESL teachers at OPS, echoing the findings of Franson (1999) and of Ruiz-de-Velasco and Fix (2000), and contrasting with the integrated content and language teaching studied by Huang and Morgan (2003), Hammond (2006), and Schleppegrell and de Oliveira (2006). Ross, Eric, and Anton all expressed a desire for school-level guidance in the decisions they were called on to make when assessing ESL students' work.

### **Conclusions**

This study explored a number of issues faced by mainstream teachers who assess ESL students' written work. Its focus on teachers' assessment practices in one Ontario high school provides a window for understanding the complex assessment decision-making undertaken by these teachers in their day-to-day teaching. It should be mentioned that our research was conducted before the release in Ontario of *Many Roots/Many Voices: Supporting English Language Learners in Every Classroom* (Ministry of Education, 2005), a document for teachers of ESL students in mainstream classes. The participants at OPS outlined their assessment decisions and dilemmas without the benefit of the advice offered by the *Many Roots/Many Voices* document.

Some limitations of this study should be noted, including the fact that the participants' assessment practices were self-reported and were elicited through one-time interviews. The number of participants was small; in addition, the research setting was a relatively privileged private school that offered small classes, tutoring opportunities, and focused preparation for postsecondary education for all its students. Of note also is that the inter-

views prompted the participants to discuss their assessment approaches in terms of two broad categories of students, namely, ESL and non-ESL.

Nevertheless, the findings of this study are significant in a number of ways. First, we found that there were discernible differences in how all but one of the participants in this study reported assessing the written work of ESL versus non-ESL students, suggesting a need for further investigations into the role played in teachers' assessment-related decisions by the first-language status of students. The mix of achievement and nonachievement factors considered in assessment decisions varied by participant, but also in some cases according to students' first-language status: some of the participants gave greater or different weights to nonachievement factors for ESL versus non-ESL students in the same class. Although awarding marks based on a mixture of achievement and nonachievement factors is well documented in the research literature, parents and students might be confused by reporting methods in Ontario, where the provincial report card separates achievement from nonachievement factors. Confusion about what marks represent might be especially evident among ESL students and their parents if they have come from school systems based on external examinations and a perceived focus entirely on achievement.

A second area in which this study is significant relates to the professional development needs of mainstream teachers of ESL students, an issue that was highlighted by the Ministry of Education's release of the *Many Roots/Many Voices* document in 2005. If the within-school variation in assessment approaches adopted at OPS both generally and more particularly with respect to ESL students' work is found to occur in other settings, this would suggest a need for formal professional development discussions and activities for teachers of mainstream classes. The participants at OPS were committed teachers who drew on many experiences when assessing the written work submitted by their ESL students; yet they made their assessment decisions in relative isolation and in many instances on a case-by-case basis. Whether or to what extent the *Many Roots/Many Voices* support document assists or influences mainstream teachers' assessment decisions about ESL students' work might form the basis of future research.

As noted above, it is the position of the Ontario Ministry of Education (2000, 2005) that teachers of all subjects should help ESL students develop their language skills; assessing ESL students' work and providing them with feedback is arguably integral to the language learning process. Because some of the participants in our study viewed language acquisition as a rather mysterious process that was the domain of ESL-specialist teachers, however, there may be challenges associated with implementing the notion that teachers of all subjects share responsibility for fostering growth in English-language skills. Implementation efforts may need to go beyond policy mandates and support documents, incorporating collaboration between teachers of

mainstream classes and ESL-specialist staff in designing assessment and instructional strategies. In this respect, the integration of content and language teaching as outlined by Hammond (2006) and by Schleppegrell and de Oliveira (2006) may be helpful models (see also Davison, 2006). Further professional development opportunities for teachers of mainstream classes such as those suggested by Huang and Morgan (2003) may promote the integration of instruction and assessment through raising teachers' awareness of the linguistic features and demands of the subject disciplines in which high school teachers work. Without such structured professional development opportunities for teachers, however, a troubling question remains: just how ESL students' language learning needs are met in schools where support for ESL students is limited or indeed where there are no ESL-specialist teachers on staff.

### *Acknowledgments*

The authors thank the *TESL Canada Journal* Editor and reviewers for their helpful comments.

### *The Authors*

Terry Milnes is a doctoral candidate in the Faculty of Education, Queen's University. Her research interests include language policy and planning and teachers' practices relating to the instruction and assessment of second-language learners.

Liyang Cheng (PhD) is an associate professor and a director of the Assessment and Evaluation Group, Queen's University

### *References*

- Airasian, P.W., & Jones, A.M. (1993). The teacher as applied measurer: Realities of classroom measurement and assessment. *Applied Measurement in Education*, 6, 241-254.
- Brookhart, S.M. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education*, 7, 279-301.
- Cheng, L., Rogers, T., & Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: Purposes, methods and procedures. *Language Testing*, 21, 360-389.
- Cheng, L., Rogers, T., & Wang, X. (2008). Assessment purposes and procedures in ESL/EFL classrooms. *Assessment and Evaluation in Higher Education*, 33(1), 9-32.
- Citizenship and Immigration Canada. (n.d.). *Immigration overview: Facts and figures 2002*. Retrieved March 23, 2004, from: [http://www.cic.gc.ca/english/pub/facts2002/immigration/immigration\\_9.html](http://www.cic.gc.ca/english/pub/facts2002/immigration/immigration_9.html)
- Cizek, G.J., Fitzgerald, S.M., & Rachor, R.E. (1996). Teachers' assessment practices: Preparation, isolation, and the kitchen sink. *Educational Assessment*, 3, 159-179.
- Cummins, J. (1996). *Negotiating identities: Education for empowerment in a diverse society*. Los Angeles, CA: California Association for Bilingual Education.
- Davison, C. (2004). The contradictory culture of teacher-based assessment: ESL teacher assessment practices in Australian and Hong Kong secondary schools. *Language Testing*, 21, 305-334.
- Davison, C. (2006). Collaboration between ESL and content teachers: How do we know when we are doing it right? *International Journal of Bilingual Education and Bilingualism*, 9, 454-475.
- Erdösy, U. (2007). Responding to student writing in a content-area undergraduate course: Implications for ESL pedagogy. *TESL Contact*, 33(2), 28-42.

- Franson, C. (1999). Mainstreaming learners of English as an additional language: The class teacher's perspective. *Language, Culture and Curriculum*, 12(1), 59-71.
- Hammond, J. (2006). High challenge, high support: Integrating language and content instruction for diverse learners in an English literature classroom. *Journal of English for Academic Purposes*, 5, 269-283.
- Huang, J., & Morgan, G. (2003). A functional approach to evaluating content knowledge and language development in ESL students' science classification texts. *International Journal of Applied Linguistics*, 13, 234-262.
- Klesmer, H. (1994). Assessment and teacher perceptions of ESL student achievement. *English Quarterly*, 26(3), 8-11.
- Leung, C., & Mohan, B. (2004). Teacher formative assessment and talk in classroom contexts: Assessment as discourse and assessment of discourse. *Language Testing*, 21, 360-389.
- McMillan, J.H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice*, 20(1), 20-32.
- McMillan, J.H., Myran, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *Journal of Educational Research*, 95, 203-213.
- McMillan, J.H., & Nash, S. (2000). *Teacher classroom assessment and grading practices decision making*. Richmond, VA: Metropolitan Educational Research Consortium. (ERIC Document Reproduction Service No. ED447195)
- McMillan, J.H., & Schumacher, S. (2001). *Research in education: A conceptual introduction*. New York: Addison Wesley Longman.
- Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York: Macmillan.
- Ministry of Education. (2000). *The Ontario curriculum, grades 9 to 12: Program planning and assessment*. Toronto, ON: Queen's Printer for Ontario.
- Ministry of Education. (2005). *Many roots/many voices: Supporting English language learners in every classroom*. Toronto, ON: Queen's Printer for Ontario.
- Ministry of Education, Ministry of Training, Colleges and Universities. (2005). *Ontario provincial report card*. Retrieved June 9, 2005, from: <http://www.edu.gov.on.ca/eng/document/forms/report/1998/report98.html#sec>
- O'Byrne, B. (2001). Needed: A compass to navigate the multilingual English classroom. *Journal of Adolescent and Adult Literacy*, 44, 440-449.
- Ontario Universities' Application Centre. (2005). *Companion booklet: Compass. 101*. Retrieved June 25, 2005, from: [http://www.ouac.on.ca/101/pdf/b\\_compass101\\_e.pdf](http://www.ouac.on.ca/101/pdf/b_compass101_e.pdf)
- Patton, M. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Penfield, J. (1987). ESL: The regular classroom teacher's perspective. *TESOL Quarterly*, 21, 21-39.
- Rea-Dickins, P. (2004). Understanding teachers as agents of assessment. *Language Testing*, 21, 249-258.
- Ruiz-de-Velasco, J., & Fix, M. (2000). *Overlooked and underserved: Immigrant students in U.S. secondary schools*. Washington, DC: Urban Institute. (ERIC Document Reproduction Service ED449275)
- Schleppegrell, M., & de Oliveira, L.C. (2006). An integrated language and content approach for history teachers. *Journal of English for Academic Purposes*, 5, 254-268.
- Stiggins, R.J. (2001). The unfulfilled promise of classroom assessment. *Educational Measurement: Issues and Practice*, 20(3), 5-15.

### *Appendix: Interview Guide*

1. Please tell me how long you've worked as a teacher and about the types of assessment activities you use in your classes.
2. In the subject you teach, what do you consider to be the hallmarks or characteristics of good writing by your students?
3. How does the written work of your ESL students compare generally to the written work submitted by your non-ESL students?
4. How do you weight style versus content marks for ESL students when you assess their written work? How does this compare to the weighting you use for non-ESL students in your classes?
5. How important a consideration is student effort when you are assessing the written work of ESL and non-ESL students in your classes?
6. How significant a role do factors such as attendance, homework completion and class participation play in your assessment process? How would you describe similarities and differences between ESL and non-ESL students in these respects?
7. Please describe some of the experiences you've had this year with respect to assessing the written work of ESL students in your classes.