

Chinese Students' Perceptions of the Value of Test Preparation Courses for the TOEFL iBT: Merit, Worth, and Significance

Jia Ma & Liying Cheng

Test preparation for high-stakes English language tests has received increasing research attention in the language assessment field; however, little is known about what aspects of test preparation students attend to and value. In this study, we considered the perspectives of 12 Chinese students who were enrolled in various academic programs in a Canadian university, and we examined their perceptions of the value of the TOEFL iBT test preparation courses they had taken prior to admission to that university. Drawing on Scriven's (1998, 2007) model, which conceptualizes value as merit, worth, and significance, the thematic findings that emerged from the interviews were triangulated with document analysis and revealed that the value of TOEFL iBT test preparation courses was attributed to (a) quality (merit), including teachers, instruction, and instructional content; (b) benefit (worth), such as efficiency and reorientation to the TOEFL iBT and English skills development; and (c) engagement and affective attainment (significance), in learning communities and sources of motivation. However, the perceived value varied among these students, and the factors related to such variability included both internal factors (e.g., students' motivations for taking test preparation courses) and external factors (e.g., limited time to prepare for the TOEFL iBT). Implications of the benefits are discussed within the context of Chinese students in the Canadian academic milieu.

La recherche dans le domaine de l'évaluation linguistique se penche de plus en plus sur la préparation aux examens d'anglais à enjeux élevés. On connaît mal, par contre, quels aspects de la préparation aux examens les étudiants trouvent importants et sur lesquels ils se concentrent. Dans cette étude, nous avons considéré les perspectives de 12 étudiants chinois inscrits à divers programmes académiques dans une université canadienne et avons examiné leurs perceptions de la valeur des cours de préparation au TOEFL iBT qu'ils avaient passé avant d'être admis à l'université. Le modèle de Scriven (1998, 2007), qui conçoit la valeur comme étant le mérite et la signification, a servi d'inspiration. Les thèmes qui sont ressortis des entrevues ont été validés à partir d'une analyse documentaire; les résultats indiquent que la valeur des cours préparatoires au TOEFL iBT était attribuée à (a) la qualité (le mérite), y compris les enseignants, l'instruction et le contenu instructif; (b) les avantages, tels l'efficacité et la réorientation vers le développement d'habiletés relatives à la langue anglaise et au TOEFL iBT; et (c) l'implication et la réalisation affective (signification) dans les communautés

d'apprentissage et les sources de motivation. Toutefois, la valeur perçue variait d'un étudiant à l'autre, et les facteurs liés à cette variabilité incluaient des facteurs internes (la motivation de suivre des cours préparatoires, par exemple) et des facteurs externes (contraintes de temps quant à la préparation aux examens). Nous en discutons les bienfaits et les répercussions dans le cas d'étudiants chinois dans le milieu académique canadien.

Recent statistics show that China has ranked as one of the top source countries of international students in Canada (Canada Immigration and Citizenship, 2015). For these Chinese students, English language tests are the gatekeepers for entrance into academic programs in Canada. Scores on these tests, along with other academic requirements, are used as criteria to determine admission into Canadian academic programs. Because of these high stakes, obtaining satisfactory scores on English language tests is the first hurdle Chinese students must jump in their attempt to gain admission. Therefore, many prospective Chinese students take test preparation courses in order to obtain the scores necessary to meet language requirements (e.g., Alderson & Hamp-Lyons, 1996; Doe & Fox, 2011). In fact, test preparation for high-stakes English language tests has become a prevalent social and educational phenomenon in China (Ma, 2013) and, after starting sometime between the late 1980s and early 1990s, has become a fast-growing, profitable industry (Matoush & Fu, 2012; Wang, 2007; Xu, 2007). Consider, for example, New Oriental, a leading private educational institution in test preparation in China. It is reported that 70% of Chinese students in universities in the United States and Canada were trained at this institution prior to their overseas studies (Tang, 2010). New Oriental is only one of an enormous number of test preparation centres in China.

Researchers have pointed out that cut scores of English language tests for admission purposes are intended to predict test-takers' academic success in studies in English rather than their ability to fully function without support in English in an academic context (e.g., Zareva, 2005). In fact, test preparation practices that some test-takers are exposed to mostly mimic test-like items/tasks, which may be different from tasks in a real academic context (e.g., Fox & Curtis, 2009). It is not surprising, therefore, to observe that nonnative English-speaking students may enter academic programs with the required English test scores but encounter difficulty in communicating in English (e.g., Fox, 2005; Read, 2008). Although many empirical and pedagogical efforts have been made to understand and support nonnative English students' adjustment into their academic programs (e.g., Cheng & Fox, 2008; Ranta & Meckelborg, 2013), there is limited understanding of what these students experienced when they prepared for English language tests and/or English academic study.

Given the large and increasing proportion of Chinese international students in Canadian educational institutions and the great popularity of test

preparation courses among Chinese students, it is worthwhile exploring their experiences in English language test preparation courses before their entry into English academic programs to determine which aspects of these courses were effective or ineffective, and which factors influenced their perceived effectiveness. The students' perspectives about how preparatory courses are valued, or what the students attended to, may benefit English language instructors working with Chinese students on Canadian university and college campuses to understand their preferences for learning English.

This study thus focused on the Chinese students who were successfully admitted into Canadian educational institutions with prior test preparation experience and with the required TOEFL iBT (Test of English as a Foreign Language Internet-Based Test) scores. Drawing on Scriven's (1998, 2007) conceptualization of value, this study examined the following research questions:

1. How do Chinese students judge the quality of the TOEFL iBT test preparation courses they completed?
2. Which aspects of the test preparation courses do Chinese students consider most beneficial?
3. What importance do Chinese students associate with TOEFL iBT test preparation courses?

Literature Review

Previous Research on Test Preparation from Students' Perspectives

Although test preparation has been researched from various perspectives, the common ground that these empirical studies have is similar—how and to what extent test preparation can influence teaching and test performance. The difference in test preparation effects from previous studies (e.g., Montgomery & Lilly, 2012) suggests the need for further empirical exploration, especially from students' perspectives, to understand why test preparation may or may not improve test performance.

Previous studies on test preparation have focused on students' traits (e.g., background characteristics and expectations) and the process of participation in test preparation (e.g., Liu, 2014; Mickan & Motteram, 2009; Yu, 2012). Researchers have found that students who have not received regular instruction in English and students with lower language proficiency are more likely to take test preparation courses (e.g., Cole, 1982; Ferman, 2004; Gan, 2009). In addition to preparing for a particular test, motivation for taking test preparation courses includes the desire to learn academic skills, to reduce anxiety, and to meet the requirements and expectations of parents or employers (e.g., Green, 2006a; Yu, 2012); the test-taking strategies that students acquired and practiced in test preparation can also be related to their motivation for taking the test (Doe & Fox, 2011).

Studies focusing on Chinese students' test preparation have revealed similar findings. The instruction Chinese students received from test preparation courses have included familiarization with test requirements and test formats, the application of test-taking strategies (e.g., Chau, 2008, Yu, 2012), and strategies for improving particular English skills (e.g., Liu, 2014). Taking test preparation courses has positively correlated with Chinese students' test scores (Liu, 2014), but certain preparatory strategies (e.g., writing in a standardized way for all writing tasks) might not contribute to improving English skills (e.g., Chau, 2008). The findings of these studies indicate that test preparation courses, depending on the nature of these courses, may or may not enhance Chinese students' outcomes on English language tests.

However, most of these previous studies were conducted from the teachers' perspectives; with regard to actual English language learning outcomes, there is limited understanding of how students perceive test preparation experience, specifically "about which aspects of a class are attended to by the learners, what they learn from them, whether the attention given to the test is rewarded" (Green, 2006b, p. 364). Evidence from students' views, such as which test preparation practices they thought worked and which did not, might help to provide important evidence to explain why test preparation practices do or do not contribute to better test performance, thus better predicting academic success after admission. Except for some quantitative evidence, based on survey data, that approximately 80% of Chinese students thought test preparation courses for the TOEFL iBT were "useful" or "absolutely essential" (e.g., Yu, 2012), there is limited evidence detailing aspects of test preparation courses that are actually "attended to" by students, or students' perceptions of the value of test preparation courses.

Conceptualization of Value

Drawing on program evaluation and evaluation research that centre on judgements of value, this study used Scriven's (1998, 2007) conceptualization of value as a suitable model to unpack the concept of value. *Value* is "the root term of evaluation" (Lincoln & Guba, 1980, p. 61), and the goal and function of evaluation is to judge the value of something being evaluated (i.e., an evaluand). As Scriven (1998, 2007) has argued, evaluation is the systematic and objective process of determining the merit, worth, or significance of an entity. The properties of merit, worth, and significance have thus been used as a conceptual logic model to determine the value of an entity. Researchers in the field of evaluation theory have defined the three attributes of merit, worth, and significance to make distinctions and also to facilitate the application of this conceptual model (e.g., Lincoln & Guba, 1980; Scriven, 1998, 2007). *Merit* comprises the intrinsic characteristics of the entity being evalu-

ated and is equivalent to inherent quality regardless of the context in which the entity is situated. *Worth* is the extrinsic property of the entity being evaluated, and it is equivalent to cost-effectiveness. Unlike merit, an inherent quality, worth is a context-dependent attribute, and the judgement of worth depends on the interaction between the entity and the local context. *Significance* is the meaning attached to the entity; thus, it is equivalent to importance.

It can be seen from these definitions that the conceptual model of merit, worth, and significance clearly address the different layers of value and also include the consideration of interaction between the entity and the context. As Lincoln and Guba (1980) stated, because worth is a context-dependent attribute, the judgement of worth must be based on "an ad hoc evaluation" in the context in which the entity is used. That is to say, the attribute of worth should be determined after the entity being evaluated has been used in a context. Because the participants in the current study are Chinese students who took test preparation courses and achieved scores on English language tests sufficient to be admitted into academic programs, this conceptual model is suitable; their judgement of value constitutes "an ad hoc evaluation." Students' perceptions of value can be analyzed using this model to include both the evaluand of test preparation course and the contexts related to test preparation courses. Therefore, by mapping them onto Scriven's conceptual model, the research focus of Chinese students' perceived value can be examined through three aspects: (a) the quality of TOEFL iBT test preparation courses Chinese students took, (b) the cost and benefits of test preparation courses, and (c) the importance, if any, the students associate with taking test preparation courses. These aspects based on the conceptual model of value help frame the research questions of this study.

Method

Participants

A university in the province of Ontario, Canada, which has a steady enrollment of international students from China (Chinese students), was the context of this study. The participants recruited were students in either undergraduate or graduate programs who had enrolled in preparation courses in China prior to taking the TOEFL iBT. These criteria were used to ensure that the participants were able to provide rich information regarding their experiences in their test preparation courses.

Participant recruitment was conducted through two streams: we first sent invitation e-mails through the electronic mailing list of the Chinese Students and Scholars Association at this university, and then used the snowball technique of asking Chinese students who responded to the initial invitation to nominate other potential participants. Three Chinese students indicated their

willingness to participate. Using the snowballing technique, 9 additional students subsequently agreed to participate. Altogether, the participants were 12 Chinese students enrolled in a variety of majors, 8 in graduate programs (see Table 1). All of them had studied English for more than 10 years in China. Six students reported that their TOEFL iBT scores ranged from 96 to 105, and the rest indicated that they had passed the threshold score of 88 set by the university. Two students in PhD programs were planning to enroll in an English for Academic Purposes (EAP) program because it was required for international students assigned to teaching assistantships.

Table 1
Profiles of Students' Demographic Information

	Mean or counts	Percentage
Age	<i>M</i> = 24	
Gender		
Female	4	33
Male	8	67
Academic program		
Undergraduate	4	33
Graduate	8	67
Major		
Science and engineer	6	50
Social science	2	17
Medical education	1	8
Businesses and management	3	25
English study in China (years)	<i>M</i> = > 10	
TOEFL iBT score	<i>M</i> = 93	
Planning to register in EAP programs	2	17

As shown in Table 2, 9 of the 12 participants wrote the TOEFL iBT more than once to achieve a satisfactory score for application. Six students chose intensive test preparation courses of two to three weeks in length, while the others chose extended courses ranging from 1 month to 2–4 months in duration. Most students received 60–80 hours of instruction. A class size of more than 100 students is typical of Chinese-style test preparation (Xu, 2007). The courses at branches of the same, national-chain, English training centre were preferred by a majority of students, although local centres were also chosen because of students' preference for small class sizes.

Table 2
Students' Background Characteristics by Their Test Preparation Experience ($n = 12$)

			Count	Percentage
Number of times the TOEFL iBT was taken	1		2	17
	2		7	58
	3		2	17
	Missing		1	8
Features of test preparation courses	Duration	< 3 weeks	6	50
		1 month	3	25
		2–4 months	3	25
	Instruction hours	< 20 hours	1	8
		60–80 hours	8	67
		> 100 hours	3	25
	Class size (Students/class)	< 30	5	42
		30–99	4	33
		> 200	3	25
Schools	National chain	10	83	
	Local	2	17	

Data Collection

Data collection was carried out using semistructured interviews, document collection, and follow-up interviews with all 12 students. From October 2013 to February 2014, a one-on-one interview with each student was conducted, and each interview was about an hour in length. All the interviews were audio recorded and followed an interview protocol (see Appendix). At the end of the interviews, the students were asked to provide any instructional materials they had obtained from test preparation courses, if available; these collected materials were used for document analysis. Five students provided the electronic documents they kept on their laptops, including class notes, in-class exercises, *Ji-jing*¹ (the test-takers' reflections of the test-taking experience and test items), and posts on TOEFL iBT test preparation websites/online forums or on teachers' blogs/social media. These documents were collected for data triangulation with the interview data. After transcribing all interview records and reviewing the transcripts and documents, 15-minute follow-up interviews (8 face-to-face and 4 by telephone) were conducted with the students, as a member-checking mechanism, if their accounts in transcripts or the documents they provided needed to be clarified or explained.

The interview protocol had four sections. Section 1 focused on demographic information (e.g., gender, programs, major, number of times the TOEFL iBT had been taken) and information regarding their test preparation experience (e.g., format of the test preparation courses, duration and hours of these courses). Sections 2 to 4 were designed according to Scriven's (1998, 2007) conceptual model of value, and each of these three sections included open-ended questions to elicit students' opinions of the merit, worth, and significance of the test preparation courses. The interview protocol was sent to a researcher with expertise in evaluation theory to ensure the interview questions in Sections 2–4 were phrased accurately to reflect the definitions of merit, worth, and significance in Scriven's conceptual model of value. This process allowed us to evaluate the internal validity of this interview protocol. Following this researcher's recommendations, the interview questions on worth were rephrased to allow participants to think about the worth by comparing alternative approaches of test preparation (e.g., taking test preparation courses vs. self-test preparation).

Data Analysis

Because of the qualitative nature of the interview data, both deductive (theory-driven) and inductive (data-driven) approaches were employed (e.g., Patton, 2002). The data analyses were conducted sequentially, using a deductive approach followed by an inductive approach. The deductive approach was employed to categorize interview transcript data into three categories—merit, worth, and significance. This approach was appropriate for generating overarching categories to understand different layers of these students' perceived values.

The inductive approach was used in the second round analysis because Scriven's conceptual model of value provides only an overarching level of constructs (i.e., merit, worth, and significance), and there was no enabling subordinate conceptual structure to be used to further analyze the data within each category. The inductive or data-driven approach to exploring qualitative data is often used to allow qualitative researchers to work from specific, meaningful units of transcripts and/or observations to broader categorizations (Corbin & Strauss, 2008). We first identified the meaningful segments of students' opinions, then grouped the segments addressing the same issue into one code, and finally labelled the code to represent the core meaning (Saldana, 2009). This process was repeated until saturation was reached. All of the codes were analyzed and grouped into themes according to the meanings they represented and the relationship among these codes (Saldana, 2009).

The documents were then reviewed by mapping them onto the categories and themes generated from the interview data. For example, Wilson, a PhD student in mechanical engineering, provided two documents that his peers shared—*Ji-jing* on TOEFL iBT test preparation forums—and he thought these documents were very helpful for understanding real test-taking pro-

cedures. These documents were then used as triangulation of the theme of learning communities generated from the interview data. All the collected documents were reviewed in this way to triangulate themes and categories and to strengthen the credibility of research findings (e.g., Bowen, 2009).

Results

Three categories and six themes (see Table 3) were generated from the interview data and were cross-referenced with collected documents: (a) *merit*—teachers, and instruction and instructional content; (b) *worth*—efficiency of test preparation, and reorientation to the TOEFL iBT and English skills development; and (c) *significance*—learning communities, and sources of motivation and inspiration. Categories and themes are elaborated to demonstrate how each represented the value of test preparation courses.

Merit

The students considered the merit, or quality, of TOEFL iBT test preparation courses from two perspectives: (a) teachers and (b) instruction and instructional content.

Teachers. The students emphasized two characteristics of their teachers: overseas educational background and TOEFL iBT test-taking background. These characteristics were important determinants when these students judged the quality of their test preparation courses. The most impressive teachers named by many students shared one common characteristic—overseas educational background—because these teachers had first-hand experience of the educational contexts that these students also hoped to experience. These teachers used their own examples to illustrate that the TOEFL iBT score was important to gain entrance into academic programs and that adequate English skills were fundamental to complete academic programs. Similarly, some teachers' real-time TOEFL iBT test-taking experience was valued because of many practical suggestions they were able to provide. Jack, a Master's student in physics, commented on a suggestion his teacher gave to deal with a possible situation in the TOEFL iBT speaking section:

The speaking section requires both listening and recording of your answers. The speaking [module] my teacher mentions, if the test-taker beside you started the speaking [section] earlier, there was the risk that you might listen to the lecture while s/he had started talking. This will definitely interfere your listening. So he suggested saving some time in the reading section by skipping the screens of test directions. Even though the saved time could not help start the speaking section much earlier, you might not be so left behind [to start speaking section]. This is a very important detail from my teacher's real-time test taking.

Table 3
Matrix of Thematic Findings of the Perceived Value
of TOEFL iBT Test Preparation Courses

		Source of data	
		I	D
Merit			
Theme One	<i>Teachers</i>		
Code 1	Overseas educational background	×	
Code 2	TOEFL iBT test-taking background	×	
Theme Two	<i>Instruction and instructional content</i>		
Code 3	Structured instruction	×	×
Code 4	Reliable instructional content	×	×
Worth			
Theme Three	<i>Efficiency of test preparation</i>		
Code 5	Time and effort efficiency	×	
Theme Four	<i>Reorientation to the TOEFL iBT and English skills development</i>		
Code 6	Rethinking the TOEFL iBT	×	
Code 7	Reorientation to English skills development	×	×
Significance			
Theme Five	<i>Learning communities</i>		
Code 8	Forming communities	×	
Code 9	Participating in communities	×	×
Theme Six	<i>Sources of motivation and inspiration</i>		
Code 10	From peers	×	
Code 11	From teachers	×	

Note. I = interviews; D = documents.

It was noted that both overseas educational background and TOEFL iBT test-taking background were part of the teachers' real-life experiences of situations that their students planned to experience. Therefore, the experiential suggestions for taking the test and practicing English skills were attended to because, as Clark, a PhD student in electrical engineering, said, "these teachers, who have their own experience in real situations, know what is needed; and compared to the teachers strictly following books, I like and trust the teachers with experiential pedagogy." However, the availability of teachers with overseas educational background or TOEFL iBT test-taking background was limited.

Instruction and instructional content. The structure of the pedagogy and the instructional content were also identified as determinants of the quality of test preparation courses. All the students reported similar course structure: four modules of reading, listening, speaking, and writing that aligned with four test sections of the TOEFL iBT. With limited time allocated for TOEFL iBT test preparation, this structure provided “a starting position to assign and break down the holistic test preparation plan into four major parts,” as Rachel, an undergraduate student in math and economics, commented.

It was noted that most of the students (83%) took test preparation courses from branches of the same national chain training centre. The reason for this choice was that this centre was perceived as professional and authoritative in terms of its instructional content. The sources of the instructional content, according to the students’ accounts and their documentation of in-class exercises, were mostly from the official website of the Educational Testing Service (ETS), the Official Guide, or the Test Practice Online (TPO) published by the ETS. Justin, a Master’s student in business, commented,

Compared to the segmented pieces of information you put together to map out the picture of the TOEFL iBT, the instructional content and information presented to you by this educational institution is more reliable, and you tend to have more trust, ... because they are professionals [in the test preparation business].

The instructional content that these students described included sequential illustrations of test features (e.g., test formats, scoring criteria, task content/topics, designated time), demonstration of test-taking strategies, management of affect (e.g., anxiety, confidence), strategies for developing English skills, and/or tips on the test-taking process. However, the coverage and the proportion of various aspects of the instructional content varied across teachers. For example, some teachers explained the importance of content knowledge in comprehending reading texts and covered some discussion of social values (e.g., gender equity, environment awareness) in their instruction. Some students thought such discussion and understanding helped them to answer relevant test items. However, other students showed negative attitudes toward such instructional content. Maggie, a Master’s student in chemistry, recalled the teacher of a reading module who extended a discussion from a reading passage about evolution to presidential elections in the United States; Maggie had no idea of how this discussion was related to test preparation, so she thought the discussion was “totally waste of my time.”

Worth

When judging the worth of their test preparation courses, the students took the perspective of examining the cost and benefit. Two themes emerged: (a) efficiency of test preparation and (b) reorientation to the TOEFL iBT and English skill development.

Efficiency of test preparation. All the students indicated that they would take test preparation courses again if they were given one more chance to prepare for the TOEFL iBT, because this was the most time-efficient method to prepare for the TOEFL. They believed that, if not taking preparation courses, they would otherwise either take much more time to familiarize themselves with the test or take the test more times to achieve a satisfactory score. The time-cost equation was what they considered and balanced in the overall process of academic applications, and taking the TOEFL iBT was only one of many requirements. Clark made a clear comment on time efficiency:

I did a cost calculation. If I had not taken the TOEFL iBT test preparation course and if I had not got a satisfactory score, I would have probably had to wait till next year to apply [for academic program], and the cost would be far more than the time and money I would spend now. In this case, I think it [taking test preparation courses] is worthwhile.

Some students related the worth to efficiency of effort and thought that the structured teaching was helpful to regulate their efforts. These participants commented on lack of self-control and self-efficacy as challenges to preparing the TOEFL iBT on their own, because they might, without the test preparation course, be disorganized and inefficient in terms of time management. Mary, a PhD student in geography, said, "I did not want to look at it [the TOEFL iBT] at all if nobody guides me how to," and the intensive course was very useful because she knew "what aspects to direct efforts onto" by following structured teaching.

Reorientation to the TOEFL iBT and English skills development. Many students stated that they had heard from many test-takers who had achieved high scores with the help of test-taking strategies. Naturally, learning test-taking strategies related to the TOEFL iBT was the main purpose for most students taking preparatory courses. However, some students recalled that their teachers, especially those with overseas educational experience, emphasized that the TOEFL iBT evaluates English language proficiency in an academic environment and repeatedly advised students not to rely solely on test-taking strategies to improve the TOEFL iBT score, as that would not increase their score substantially. Some students recalled the examples their teachers presented to help them understand that scores were based on the English skills demonstrated through task performance on the test. Jenny, a Bachelor's student in nursing, stated that her teacher gave detailed instructions on a scored writing sample in the Official Guide and explained why this sample was rated highly by mapping it onto the scoring criteria:

The [teacher's] instructions made me understand that the clear logic throughout this writing sample ensures the coherence of the ideas presented, and one portion of the high score was awarded to its clear

logic; and the high score is definitely associated with the skills presented in this writing [not test-taking strategies].

This reorientation to the constructs of the TOEFL iBT, or the recognition that high scores would not be within arm's reach by only employing certain test-taking strategies, was an unexpected benefit to some students. The shared understanding of the TOEFL iBT as a proficiency-based test guided students to attend to the instruction on developing English skills. Some students reported that they practiced English skills outside of class time and/or after the preparation course was over. The limited class time restricted teachers to the introduction of some strategies for building English skills; therefore, some students said they tried out and practiced these strategies after class. For example, Mary recalled her teacher demonstrating how to keep notes of the logic of a talk and related details with the help of a flow chart. The logical flow chart changed her previous strategy of taking note of everything she heard. When repeatedly practicing this strategy on listening tasks after class, she noticed that it worked well and "my listening is improved immediately."

In addition, some students planned a period of time (e.g., 2–3 months or longer) after completing test preparation courses as a reinforcing period; building their English skills was a common focus for reinforcement before actually taking the TOEFL iBT. Rachel increased her score from 86 to 104 in six months, and she attributed the score improvement to the continual practice of English skills, which she believed was guided by the appropriate reorientation to the constructs of the TOEFL iBT and English skills development.

However, reorientation to the TOEFL iBT and English skill development was not experienced by all of the students. Some of their accounts revealed that their attitudes toward the reorientation were influenced by both personal and contextual constraints. Some students said that the only purpose for taking test preparation courses was to familiarize themselves with the test and that they had no expectation of improving their English skills through these courses. Given this instrumental orientation, all of their attention was directed to instruction related to the test itself. When personal constraints (e.g., a narrowed attitude toward learning) were confounded with contextual constraints (e.g., a limited timeline), it was unlikely that the students would shift their orientation to the need to develop English skills. As Justin commented, "I have a very tight schedule [for preparing for the TOEFL iBT], and I do not want to improve my English proficiency through this course; I *just* need a score [emphasis in original tone], and this is most important to me."

Significance

Significance was defined as how the students perceived the importance of test preparation courses, and this importance was related to two factors: (a) learning communities and (b) sources of motivation and inspiration.

Learning communities. One advantage of test preparation courses was the opportunity to meet peers who were involved in a similar application process. Shared goals and experiences enabled discussions and the exchange of ideas among peers; therefore, communities were formed as platforms to facilitate discussions, either by students or teachers. Some communities were virtual, such as online chat groups or groups in online forums, and others were face-to-face, such as lunch-mate groups. For example, Mary recalled that she and five other girls in the same intensive course soon became a group after taking lunch break together everyday. She felt “better” being in this group because she could talk through the difficulties and stress she encountered during the test preparation process and knew that she was not “the only one who had experienced those difficulties.” However, some students admitted that the possibility of forming communities depended on interactions with peers in class, and not all the students in this study reported the emergence of learning communities in their accounts. Lindsay, a Master’s student in business management, mentioned that only after being assigned to the same group for performing an oral task did she become familiar with the students sitting close to her; they then started an online chat group. In addition, some students mentioned that their teachers used social media platforms to form communities, such as blogs, Weibo (a Chinese microblogging site akin to Twitter), and public account subscriptions on WeChat (a mobile communication application similar to WhatsApp).

Participation in learning communities mainly included sharing information and resources among peers and teachers. The students commented that they valued the resources shared in these communities for two reasons: (a) they could access useful resources that they were not aware of previously, and (b) the shared resources reviewed and recommended by peers and teachers were a helpful starting point, given the overwhelming range of resources available on the market. Wilson, a PhD student in materials engineering, who practiced listening skills with the podcast *60-Second Science* by Scientific American, said he came to know this resource from a *Ji-jing* post shared by his peers in an online chat group. He clearly remembered that this *Ji-jing* was a detailed reflection of a test-taker’s experience of preparing for the TOEFL iBT, and *60-Second Science* was the resource this test-taker had used. Wilson appreciated this shared resource because he thought it helped him improve his listening skills. The learning communities that teachers initiated—for example, an online group—were an extension of the class, and many students said they read posts related to the test and to English language learning. Clark recalled that the writing teacher posted on her blogs the transcripts of a speech that Obama delivered in Russia and highlighted some sentences in blue to focus students’ attention on sentence structure and vocabulary. Clark particularly mentioned he had constant interactions with his teachers, including this writing teacher, on social media platforms, and he admitted that the

rapport he developed with his teachers through the interactions increased his trust for them and, in turn, for their instruction.

Sources of motivation and inspiration. Many students regarded taking test preparation courses as an experience of gaining inspiration, encouragement, and psychological support from peers and teachers; attending test preparation was more like a ritual to equip them with “spiritual powers,” in William’s words. The test preparation courses attracted groups of youth with similar goals working hard together toward achieving them. Commenting on the class as a platform for encouragement, Clark noted:

I came to the class before 8 a.m. and found that four or five students had been there and started reciting vocabulary. Then I immediately knew what I should do. ... I also found that my English is not that good compared to many students in class. Those whose English were far more better worked really hard, how could I waste time? This atmosphere was the source of my motivation.

Students reported that many of their teachers shared the successful experiences of their own or of previous students. These narratives were very inspirational because those who succeeded had similar backgrounds to these students. Knowing what efforts others had made—and how they had succeeded—encouraged the students to work harder. William recalled,

Those successful stories my teachers told mentally supported me to insist on sitting in a classroom with 300 students in hot summer days; ... it sounded tough but I did not feel so because it was where I got inspired.... The mental power from these success also conveyed a message of persistence to me, which I think is so important not only for taking the TOEFL iBT but also for each step of the academic application process.

Discussion

The results of the study revealed that the students’ judgement of the value of the preparatory courses was derived from perceived quality, benefits, and importance. The results clearly indicate that these preparatory courses presented students with experiences of preparing to take the TOEFL iBT (e.g., instruction on test features) and/or of developing the English skills required for academic studies in English (e.g., reorientation to English language skills). However, these 12 students gave various accounts of their experiences in these preparatory courses (e.g., instrumental motivation for taking the preparatory courses vs. recognition for building English skills), and their experiences directly influenced how they perceived the effectiveness of each aspect of the preparatory courses, such as instruction and peer support. Therefore,

our discussion will focus on the variability and the factors contributing to the variability.

Similar to many test-takers in previous studies (e.g., Green, 2006a, 2006b), the students paid great attention to the content of the test (e.g., test structure, test formats) and how the test is administrated (e.g., test-taking process). This is what the Chinese students mostly aimed to become familiar with in their preparatory courses. This attention was related to their motivation for taking such courses, for example, to have “a starting position for the holistic test preparation plan” or “I just need a score [to meet the application requirement].” Although all the students reported similar motivation for test familiarization, it is interesting to note the shift from this instrumental motivation to an intrinsic understanding of the need to develop English language skills. From the accounts of the students who recognized the importance of and practiced English skill-building strategies, it is clear that their teachers and the pedagogy used by their teachers had a great influence on shaping this important shift.

In most students’ accounts, teachers with overseas educational experience were valued because their background of English academic study helped to construct pedagogy suitable to their students’ future academic study needs. To these students, their teachers were a direct source of information regarding the academic situations that students would face; therefore, they believed their teachers “kn[e]w what is needed [in real time settings].” It is not surprising to note that when these teachers stated that building sufficient English skills was key to success in English academic study, students tended to follow faithfully and practice the strategies taught, such as using logic flow charts to practice listening skills. Similar findings were reported in recent studies of how teachers are important determinants of test preparation courses. Green’s (2006b) study found that the learning outcomes of students in test preparation courses were related to teachers’ choice of instructional content. Gibson and Swan (2008) indicated the benefits of shared linguistic background and educational experience on teachers’ pedagogy in test preparation courses. Since teachers’ first-hand experience is valued, teachers with relevant academic or content knowledge background may be also valuable to students, and instruction in some first-hand academic tasks would be helpful for students to understand what is required in a specific major of study (e.g., writing format, structure, logic) and to provide students with first-hand experience that they can connect to their academic study.

Students also valued their teachers’ first-hand test-taking experience because they needed to know what was expected in the test-taking process. The importance of having testing researchers and practitioners working with language tests has been demonstrated in recent studies (e.g., DeLuca, Cheng, Fox, Doe, & Li, 2013; Fox & Cheng, 2015). For example, the situation regarding the testing environment (e.g., the different pace of test-takers’ completion of test sections, as Jack mentioned) would interfere with test-takers’ perfor-

mance on tasks and also on the test. Teachers' experiential recommendations for the test-taking process may help students become aware of situations that might pose barriers to demonstrating their actual English language proficiency. Instructors of English language programs would be in a stronger situation to support their students by actually having taken the test their students are required to take, or by at least familiarizing themselves thoroughly with it and the testing environment in which students take it.

Students' accounts indicated that their timeline was closely related to their perceptions of the value of test preparation courses, that is, the efficiency of test familiarization, or their conception of the need to improve English proficiency to best achieve satisfactory test scores. Some students attended 3-week intensive courses and took the TOEFL iBT right after completing the courses. These students "just need a score" and they wanted to take the test when they had the freshest ideas of what the test was about. Other students planned more time for test preparation (e.g., at least 6 months before the test), and they chose extended preparatory courses. With the longer plan, these students were more likely to practice the strategies that their teacher had introduced to build their English skills before taking the test. Rachel's score improvement from 86 to 104 indicated the effectiveness of her English-skill building practices. However, not all of the students who had a longer timeline for preparation reported substantial score increases. This may raise a question for language testing researchers and language instructors: how long do students need to improve English proficiency for this to be reflected in a higher test score?

Engaging in learning communities and becoming motivated and inspired by peers and teachers were perceived as important and useful during the test preparation process. Learning communities are a means of extending in-class instruction, which has time constraints, and students appreciated information sharing and interactions with peers and teachers when participating in such learning communities. An important part of the documents that some students had collected through communities (e.g., online forums) were narratives of high TOEFL iBT achievement as a result of the test preparation process. Many of the accounts demonstrated how students had achieved high scores starting from a moderate level of English proficiency (as indicated by a national English test score in China). Successful experiences of peers like this are sources of strength and inspiration that Chinese students value; as William said, "taking test preparation courses was more an experience of being inspired." Students obtained encouragement and inspiration from examples of students with similar experience and used this affect to support them throughout the test preparation process. Although benefits of collaborative learning (e.g., working in groups or with peers) have been recognized (e.g., Cheng & Fox, 2008), an understanding of learning communities initiated by students and/or teachers, especially the support from these communities within and beyond the class boundary, remains limited (e.g.,

Potts, 2005). It may be worthwhile including students from previous cohorts in these learning communities, where they can share their experience of difficulties and challenges in adjusting to academic environments and explain how they survived and gradually thrived in their academic studies.

Conclusion

This was an exploratory study with a relatively small sample size. Nevertheless, these 12 informants provided valuable insights, allowing us to better understand what test preparation courses they valued prior to entering a Canadian academic setting. Mapping these onto Scriven's conceptualization, these students' accounts of the value covered multiple determinants of the quality (*merit*), benefit (*worth*), and importance (*significance*) of these test preparation courses. Quality was related to teachers' characteristics and instruction, benefit was judged from efficiency and reorientation to the TOEFL iBT and English skill development, and importance included engagement in learning communities and motivation to learn. This is one of the few studies that have addressed test preparation from students' perspectives, thus filling a research gap by answering the call to examine "which aspects of a [test preparation] class are attended to by the learners, what they learn from them" (Green, 2006b, p. 364). The findings could provide pedagogical insights to instructors of EAP programs in Canada, especially those working with Chinese students, to understand these students' experiences and their perceptions of suitable and effective support in English language learning. This is important, considering the fast-growing number of Chinese students on Canadian university and college campuses.

The findings of this study highlight two important aspects of what we have learned about test preparation. On one hand, students gained a changing perception of the value and need for improvement in general English proficiency. On the other hand, they retained an instrumental perception focused on reaching the admission cut score through test preparation courses. Both orientations were influenced by the participants' awareness of the amount of time realistically required to develop improved proficiency in a second language. Timing is crucial for all international students seeking a required test score for admission prior to arriving in Canada, and acquiring further language skills to be successful in their academic studies thereafter. This raises a number of issues for EAP programs: How much time does it take to develop L2 proficiency to a desired level? How do EAP programs continue to support international students in developing proficiency after admission? What is the relationship between achieved levels of proficiency and subsequent indicators of student success?

As this was a study of an exploratory nature, we note the following limitations. The participants were 12 Chinese students with TOEFL test preparation experience at one Canadian university only. Further investigations with

students from different ethnic groups and universities who have experience taking different English language tests (e.g., IELTS, Pearson Test of English [PET], Canadian Academic English Language Assessment [CAEL]) could increase the comprehensiveness and richness of data and could be compared with the findings of this study. Studies of this nature forefront the value of understanding students' perceptions of merit, worth, and significance of test preparation, which, in turn, can be used to better support their academic studies in Canada.

Note

1 *Ji-jing* is a generic term used by Chinese test-takers to refer to test-takers' reflections of the test-taking experience. Some test takers recalled test items right after they wrote the test, such as the topics of the speaking/writing tasks, and then posted these reflections online to share with future test-takers as test preparation resources. In addition to the test items, some test-takers' *Ji-jing* also includes detailed narrative accounts of how they prepare for the test, such as strategies to practice listening tasks.

Acknowledgements

The authors acknowledge funding support from the Social Sciences and Humanities Research Council (SSHRC) of Canada Partnership Development Grant. The authors also thank all the participating Chinese students who generously took the time to discuss their experience with test preparation. Both authors sincerely thank the two anonymous reviewers and Dr. Marian J. Rossiter for their valuable feedback and constructive comments.

The Authors

Jia Ma is a PhD student in the Assessment and Evaluation Group at the Faculty of Education, Queen's University. Her research interests include consequential validity of large-scale testing and connections between large-scale testing and classroom-based assessments. She recently co-authored a chapter with Dr. Liying Cheng in *TESOL Encyclopedia of English Language Teaching*.

Liying Cheng, PhD, is Professor and Director of the Assessment and Evaluation Group at the Faculty of Education, Queen's University. Her primary research interests are the impact of large-scale testing on instruction, the relationships between assessment and instruction, and the academic and professional acculturation of international and new immigrant students, workers, and professionals to Canada.

References

- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*, 13(3), 280-297. doi:10.1177/026553229601300304
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Method*, 9(2), 27-40. doi:10.3316/QRJ0902027
- Canada Immigration and Citizenship. (2015). *Evaluation of the international student program*. Retrieved from Canada Immigration and Citizenship website, <http://www.cic.gc.ca/english/resources/evaluation/isp/2015/E3-2013-ISP.pdf>
- Chau, M. (2008). *Learning to write essays or teaching to crack tests: An empirical study investigating the washback and impact of TOEFL test-preparation in China*. Paper presented at the Language Testing Research Colloquium, Hangzhou, China.
- Cheng, L., & Fox, J. (2008). Towards a better understanding of academic acculturation: Second language students in Canadian universities. *Canadian Modern Language Review*, 65, 307-333. doi:10.3138/cmlr.65.2.307

- Cole, N. (1982). The implications of coaching for ability testing. In A. K. Wigdor & W. R. Garner (Eds.), *Ability Testing: Uses, Consequences, and Controversies* (pp. 389-414). Washington DC: National Academy Press.
- Corbin, J., & Strauss, A. (2008). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (3rd ed.). Thousand Oaks, CA: Sage.
- DeLuca, C., Cheng, L., Fox, J., Doe, C., & Li, M. (2013). Putting testing researchers to the test: An exploratory study on the TOEFL iBT. *System*, 41(3), 663-676. doi:10.1016/j.system.2013.07.010
- 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, 29-54. doi:10.3138/cmlr.67.1.029
- Ferman, I. (2004). The washback effect of an EFL national oral matriculation test to teaching and learning. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 191-210). Mahwah, NJ: Lawrence Erlbaum.
- Fox, J. (2005). Re-thinking second language (L2) admission requirements: Problems with language-residency criteria and the need for language assessment and support. *Language Assessment Quarterly*, 2(2), 85-115. doi:10.1207/s15434311laq0202_1
- Fox, J., & Curtis, A. (2009). International English language testing system. In L. Cheng & A. Curtis (Eds.), *English language assessment and the Chinese learner* (pp. 112-120). New York: Routledge, Taylor & Francis Group.
- Fox, J., & Cheng, L. (2015). Walk a mile in my shoes: Stakeholders accounts of testing experience with a computer-administered test. *TESL Canada Journal*, 32(9), 65-86. doi:10.18806/tesl.v32i0.1218
- Gan, Z. (2009). IELTS preparation course and student IELTS performance: A case study in Hong Kong. *RELC Journal: A Journal of Language Teaching and Research*, 40(1), 23-41. doi:10.1177/0033688208101449
- Gibson, C., & Swan, A. (2008). *The significance of sociolinguistic backgrounds of teachers of IELTS test preparation courses in selected Malaysian institutions* (IELTS Research Report, Vol. 8). Retrieved from https://www.ielts.org/~media/research-reports/ielts_rr_volume08_report2.ashx
- Green, A. (2006a). Washback for learners: Learner and teacher perspectives on IELTS preparation course expectations and outcomes. *Assessing Writing*, 11(2), 113-134. doi:10.1016/j.asw.2006.07.002
- Green, A. (2006b). Watching for washback: Observing the influence of the International English Language Testing System academic writing test in the classroom. *Language Assessment Quarterly*, 3(4), 333-368. doi:10.1080/15434300701333129
- Lincoln, Y., & Guba, E. G. (1980). The distinction between merit and worth in evaluation. *Educational Evaluation and Policy Analysis*, 2(4), 61-71. doi:10.3102/01623737002004061
- Liu, O. (2014). *Investigating the relationship between test preparation and TOEFL iBT performance* (Report No. RR-14-15). Retrieved from Educational Testing Service: <http://onlinelibrary.wiley.com/doi/10.1002/ets2.12016/epdf>
- Ma, J. (2013). *Test preparation for high-stakes English language tests in China*. *Test Preparation Research Newsletter*, Volume 5. Retrieved from Faculty of Education, Queen's University: <http://educ.queensu.ca/sites/webpublish.queensu.ca.educwww/files/files/Research/Test%20Preparation%20Newsletter%20Volume%205.pdf>
- Matoush, M. M., & Fu, D. (2012). Tests of English language as significant thresholds for college-bound Chinese and the washback of test-preparation. *Changing English*, 19, 111-121. doi:10.1080/1358684X.2012.649176
- Mickan, P., & Motteram, J. (2009). *The preparation practices of IELTS candidates: Case studies* (IELTS Research Report, Vol. 10.) Retrieved from https://www.ielts.org/~media/research-reports/ielts_rr_volume10_report5.ashx
- Montgomery, P., & Lilly, J. (2012). Systematic reviews of the effects of preparatory courses on university entrance examinations in high school-age students. *International Journal of Social Welfare*, 21(1), 3-12. doi:10.1111/j.1468-2397.2011.00812.x

- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Potts, D. (2005). Pedagogy, purpose, and the second language learner in on-line communities. *Canadian Modern Language Review*, 62(1), 137-160. doi:10.3138/cmlr.62.1.137
- Ranta, L., & Meckelborg, A. (2013). How much exposure to English do international graduate students really get? Measuring language use in a naturalistic setting. *Canadian Modern Language Review*, 69, 1-33. doi:10.3138/cmlr.987
- Read, J. (2008). Identifying academic language needs through diagnostic assessment. *Journal of English for Academic Purposes*, 7(3), 180-190. doi:10.1016/j.jeap.2013.04.001
- Saldana, J. (2009). *The coding manual for qualitative researchers*. Los Angeles, CA: Sage.
- Scriven, M. (1998). Minimalist theory: The least theory that practice requires. *American Journal of Evaluation*, 19(1), 57-70. doi: 10.1177/109821409801900105
- Scriven, M. (2007). The logic of evaluation. In H.V. Hansen, C. W. Tindale, J.A. Blair, R.H. Johnson, & D. M. Godden (Eds.), *Dissensus and the search for common ground (Proceedings of the 7th OSSA Conference)*, CD-ROM (pp. 1-16). Windsor, ON: Ontario Society for the Study of Argumentation.
- Tang, W. (2010, April 27). Minhong Yu's talk on "happy learning and healthy grow-up" in Wuxi. *The New Oriental*. Retrieved from <http://old.neworiental.org/publish/portal0/tab410/info498224.htm>
- Wang, X. (2007). From restoration to mega-expansion: higher education reform in China: Higher education reform in China: A three decade review (1976-2006). *Analytical Reports in International Education*, 1, 7-20.
- Xu, M. (2007). *How to get a slice of the action: The analysis of China's English training industry* (Unpublished master's thesis). University of Nottingham, UK.
- Yu, G. (2012). *Preparing for TOEFL iBT speaking tasks: test-takers' experiences and expectations*. Paper presented at 34th Language Testing Research Colloquium, Princeton, New Jersey.
- Zareva, A. (2005). What is new in the new TOEFL iBT 2006 test format? *Electronic Journal of Foreign Language Teaching*, 2(2), 45-57.

Appendix

Sample Interview Questions in the Interview Protocol

MERIT

1. Why did you decide to take this particular TOEFL iBT test preparation course(s)? Why did you choose this school or centre?
2. What was the training you received on the TOEFL iBT test preparation course at this school/centre?
3. How do you think about the course structure? The resources? The materials provided by the teacher or the school/textbooks?
4. How many teachers were there in your TOEFL iBT preparation course? Which teacher(s) impressed you most? In which way did the teacher(s) impress you?
5. In addition to the training and the teachers of TOEFL iBT preparation courses, what other aspects of this school or centre (e.g., resources, lectures, workshops, etc.) contributed to your TOEFL iBT preparation?

WORTH

1. What were your expectations of taking the TOEFL iBT preparation courses? Do you think your expectations were met? Why or why not?
2. What are the benefits of taking this TOEFL iBT preparation course at this school/centre (on your test-taking, English language learning, or whatever you perceived to have benefited)?
3. Compared to alternative ways available, did you think taking test preparation courses was a better or more effective way? Why or why not?
4. Given the TOEFL iBT test score you got after taking this preparation course, do you think this course is worth the time, money, and effort you spent/made?
5. If you go back in time, would you prepare for the TOEFL iBT in the same way or differently? Would you choose a different type of preparation course? Why or why not?

SIGNIFICANCE

1. In addition to the aspect of TOEFL iBT test-taking, what other importance of taking this preparation course did you see?
2. What were the impacts of taking this course on your test-taking or academic studies (e.g., attitudes)?