

## **Language Assessment Literacy as Professional Competence: The Case of Canadian Admissions Decision Makers**

Beverly Baker  
*University of Ottawa*

### **Abstract**

This language assessment literacy project involves the collaboration of assessment professionals and admissions officers at higher education institutions across Canada. Following a survey with 20 institutions (Baker, Tsushima, & Wang, 2014), workshops were held at eight institutions across the country dealing with the use of language test scores in university admissions decision making. Recordings of these workshops were analyzed within the typology of workplace knowledge developed by Eraut and his colleagues (Eraut 1994, 2000, 2004a, 2004b; McKee & Eraut, 2012). In addition, participants commented on the usefulness of the workshop materials for their work responsibilities. Results provided insights into a) the language assessment literacy (LAL) base needed for these specific users, including both propositional and procedural components, and b) the possibilities of conceptualizing LAL for these score users as a type of professional workplace competence.

### **Résumé**

Ce projet de sensibilisation en évaluation linguistique a été effectué avec la collaboration de professionnels de l'évaluation et de responsables des admissions dans les établissements d'enseignement supérieur à travers le Canada. Suite à un sondage effectué auprès de 20 institutions (Baker, Tsushima et Wang, 2014), des ateliers ont été organisés dans huit établissements à travers le pays traitant de l'utilisation des résultats des tests de langue dans la prise des décisions d'admission. Les enregistrements de ces ateliers ont été analysés dans la typologie des connaissances en milieu de travail développé par Eraut et ses collègues (Eraut, 1994, 2000, 2004a, 2004b ; McKee et Eraut, 2012). Les participants ont formulé des observations sur l'utilité du matériel d'atelier pour leurs responsabilités professionnelles. Les résultats nous aident à mieux comprendre a) les notions en évaluation de langues les plus pertinentes pour ces utilisateurs spécifiques et b) le potentiel de concevoir la « littératie en évaluation » comme étant un type de compétence professionnelle en milieu de travail.

## **Language Assessment Literacy as Professional Competence: The Case of Canadian Admissions Decision Makers**

### **Introduction**

#### **Context**

University admissions officers are unquestionably key stakeholders in language assessment. They regularly make use of language test scores, in combination with other information, in making the high stakes decision to admit applicants to their institutions. In Canadian postsecondary institutions, there are currently well over 100,000 full-time international students, representing approximately 6.5% of student enrolment (Canadian Bureau for International Education, 2014). The majority of these students are nonnative speakers of English attending English-medium institutions. In addition to international students, there are also large numbers of French Canadians and other nonnative-English-speaking Canadians choosing to pursue their postgraduate studies in English.

Many different individuals are involved in the admissions decision-making process for these applicants, and hold a variety of titles in registration, admissions, and recruitment. The term “admissions decision makers” will therefore be used here to refer to all individuals who are involved in the decision-making process, from the registrars who approve policies on language test score use, to the admissions officers and program coordinators who communicate with applicants and provide guidance on cutoff scores. These admissions decision makers are faced with a daunting task in interpreting a growing number of English proficiency measures submitted by their applicants. Therefore, a certain base in language assessment literacy for this group (Taylor, 2009, 2013) is essential. In this report, the construct of language assessment literacy (LAL) will be explored in the context of a study involving the collaboration of language assessment specialists and Canadian university admissions decision makers.

#### **Language Assessment Literacy**

Only recently has LAL emerged as a distinct area of research within language assessment. LAL is variously referred to in the literature as a knowledge base, as well as a set of skills, principles, or competencies as applied to different test stakeholder groups (see Davies, 2008). Fulcher’s (2012) definition of LAL included skills, knowledge and abilities, awareness of the theoretical basis for assessment, and awareness of “the role and impact of testing on society, institutions and individuals” (p. 125). In a special issue of *Language Testing* dedicated to language assessment literacy, Taylor (2013) discussed this definitional issue, and stated that a key area of future research must be related to an examination of the construct of LAL itself—its distinctiveness from assessment literacy more generally (if necessary), and whether it should best be conceptualized as knowledge, a skill set, a set of principles, or all three. She proposed a differential model of LAL, showing a suggested typology of assessment literacy (AL)/LAL and the differences in emphasis for varied stakeholders. The model produced for university administrators is found in Figure 1.

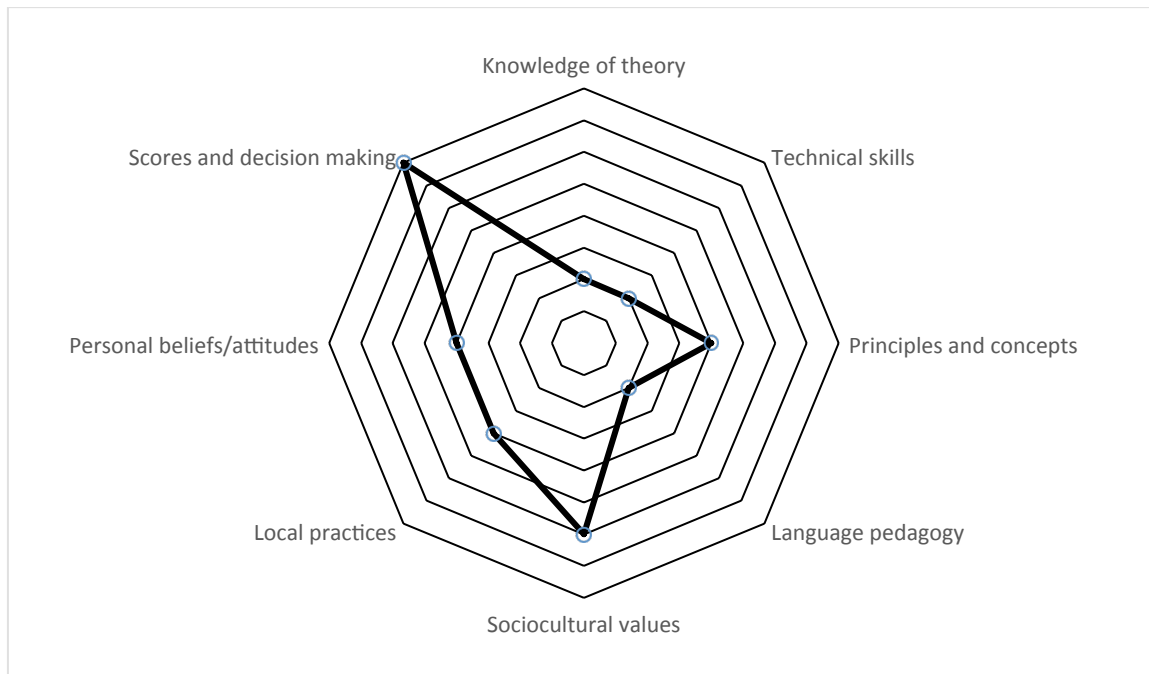


Figure 1. A model of AL/LAL for university administrators (Taylor, 2013).

There may be value in conceptualizing LAL as a *professional competency* for these test stakeholders. A competency involves acquisition of a relevant knowledge and strategic base, as well as a certain level of skill in the application of that base to situations within a community of practice.

### **Language Assessment Literacy as a Competency of Professional Decision Makers: University Admissions Officers**

Empirical LAL research with admission officers so far has focused on their reports of their knowledge about or understanding of the use and interpretation of major language tests, with most work so far focusing on the International English Language Testing System (IELTS; for a more detailed discussion, see Baker, Tsushima, & Wang, 2014). All researchers argued for a stronger base of LAL for admissions decision makers (Ginther & Elder, 2013; Hyatt & Brooks, 2009; O’Loughlin, 2011, 2013; Rea-Dickins, Kiely, & Yu, 2007). These contributions reveal the need for LAL development in admissions decisions makers. However, there is still more information needed about assessment officers’ perceptions and concerns as related to not only specific tests but to the assessment of language proficiency for admissions more generally. In addition, there is still much to be done in order to determine the exact nature of LAL required by these stakeholders, keeping in mind the wide range of language proficiency measures being used (see Baker, Tsushima, & Wang, 2014).

Determining the LAL of admissions decisions makers will require additional work to capture their knowledge of assessment concepts as related to their work, as well as their skill in applying this knowledge in the exercise of professional judgment, defined by Eraut (1994) as “the wise decision made in the light of limited evidence by an experienced professional” (p. 17). Admissions officers are different from novice teachers, or novice

language testers. They are not, for the most part, involved or interested in the development or administration of language assessments themselves. Their primary interaction with language tests is the use of scores as one source of evidence among many others to make one specific and complex decision. If professional workplace competence and processes are of interest, it would be potentially fruitful to look for insights from researchers who work on the nature and development of professional competence at work.

### **Professional Competence and Workplace Learning**

The task of establishing a working definition of a knowledge base for AL for the profession of admissions decision makers is not simplified by the opaque understandings of the knowledge base for professions in general (see Eraut, 1994). Eraut (1994) and McKee and Eraut (2012) have worked to develop an epistemology of professional competence, which is defined as follows:

... being able to perform the tasks and roles required to the expected standard. This expectation, being socially defined, will either be part of professional regulation or be determined by the micro-politics of the particular context. In either case, unlike terms like “knowledge,” “learning” and “capability,” the word competent entails a social judgment, which may vary across contexts and over time. (McKee & Eraut, 2012, p. 4)

LAL for this group might then usefully be operationalized as *competence* in the professional practice of using language test scores effectively in admissions decision making. This sociocognitive definition of competence is not incompatible with Messick’s (1989) cognitive model—more familiar to those in the area of language assessment—which viewed underlying competence as distinct from performance. In this context, competence can be interpreted as an underlying trait essential to effective job performance (i.e., effective decision making). The judgment of what is considered effective is the sociocultural aspect in this conception, specific to the situation.

Knowledge therefore represents the building blocks of competence—what professionals need to know to do their jobs well. As previously discussed, an epistemological question has arisen in the literature in LAL regarding whether knowledge, skills, attitudes, and values exist as separate constructs. Eraut (1994) used the term *knowledge* to refer to all of these constructs—both propositional and procedural knowledge, “knowing what” and “knowing how.” Process knowledge (knowing how) can be viewed as practical knowledge that may or may not be codified or explicit. Eraut (1994) discussed how moral and ethical principles are embedded in this practical knowledge, which facilitates deliberative processes such as decision making and problem solving. Eraut (1994, p. 43) divided propositional knowledge into the following three categories:

- discipline-based theories and concepts, derived from bodies of coherent, systematic knowledge (*Wissenschaft*);
- generalizations and practical principles in the applied field of professional action; and
- specific propositions about particular cases, decisions and actions.

Studies in LAL up until this point have been concentrated on which elements of the first category are necessary to stakeholders not directly involved in language assessment as a discipline, and how elements from this first category can be transformed into the second category in a way that is comprehensible and accessible. Of the third category, little has been said because language assessment researchers are not the ones to say it—specific work-related cases must be brought by the stakeholders themselves.

After a series of projects examining the professional learning trajectories of early career professionals, Eraut and his colleagues (Eraut, 1994, 2000, 2004a, 2004b; McKee & Eraut, 2012) developed a typology of workplace knowledge which was intended to be a starting point for locally specific application—“a heuristic for use in research and consultancy that reminds people of possible aspects of learning in their own context” (Eraut, 2004a, p. 265). The eight elements of the typology, referred to as a “progression typology” (Eraut, 2004a, p. 266), are briefly summarized here:

- task performance (i.e., skills generally associated with procedural memory, such as processing speed in dealing with task complexity);
- awareness and understanding (e.g., of other people, social and political contexts of the professional activity, institutional constraints)
- personal development (e.g., knowledge of the traits to be developed for professional success, learning from experience, handling emotions);
- teamwork (i.e., collaborative work);
- role performance (e.g., leadership, delegation, crisis management);
- academic knowledge and skills (e.g., research-based practice, use of evidence and argument);
- decision making and problem solving (e.g., dealing with complexity, evaluating options, sometimes under pressure); and
- judgment (e.g., setting priorities and assessing risk).

This typology is starting to be taken up as a tool in specifying workplace knowledge and learning trajectories in the professions (see Sharu, 2012, for an example within nursing). Taking this typology into account, one can observe that the great majority of the LAL work done so far with admissions decision makers has focused on academic knowledge and skills—mostly propositional, formal knowledge, or according to Taylor’s (2013) typology of LAL for university administrators, knowledge of theory as well as principles and concepts. A competency-based conceptualization of LAL that considers other elements of this typology has the potential to add substantially to this previous work, by helping to delineate propositional versus procedural knowledge, as well as explicit and tacit elements of the LAL base.

Evidence of each of these types of knowledge could conceivably be collected in a variety of ways. For example, role performance and working with others is best understood by observing admissions officers on the job or collecting artefacts from their day-to-day work, such as workplace formal and informal communications. Other types of knowledge can be obtained in face-to-face discussions, such as what was possible in the current study. However, Eraut (2004a) cautioned that collecting this information may be challenging: Often many elements are tacit; people cannot necessarily describe the nature of their own

competence. In addition, most workplace learning is informal, arising from problems and challenges or the management of change—it may not even be recognized or reported as learning.

## Method

### Research Objective

The research project reported here was an exploration of the existing LAL base and ongoing development of that base with university admissions decision makers. This exploration was undertaken with the collection of participant reports and reflections that were then classified as evidence of the various elements of Eraut's (2004a) typology of workplace learning.

The research questions for the project were as follows:

1. How can an LAL base be usefully described for users of language test scores in admissions decision making at postsecondary institutions in Canada?
2. What useful materials can be created to develop this LAL base for admissions decision makers?

To answer these questions,

- data must be collected that allow for the elaboration of a profile of the LAL competency base for these score users,
- materials need to be created that aim to build this LAL base, and
- these materials need to be provided to the decision makers for them to evaluate based on their usefulness.

“Useful” materials have to be perceived as useful by all parties, and lead to increased competence in the activity for which the scores are used. In the sociocultural view of competence discussed earlier, competence in full cannot be established only by the language testers (who can only judge whether assessment principles are understood correctly).

A previous phase of this project was a survey with admissions decision makers at 19 institutions across Canada in 2012 (see Baker, Tsushima, & Wang, 2014). The survey, adapted from O'Loughlin (2011), contained open-ended questions focusing on knowledge, beliefs, and levels of confidence in making use of language test scores in decision making. Questions were analyzed for emergent themes related to language assessment principles, as well as specific content for future workshops. Detailed results of this survey are reported in detail elsewhere (Baker, Tsushima, & Wang, 2014). Briefly, they indicated a tacit knowledge of language assessment concepts (such as the limitations of a single measure for decision making) as well as common questions and concerns regarding issues such as test fraud, the establishment of cutoff scores, and the relationship between test scores and university success. The current study reports on the second phase of the project, and involved conducting workshops that were informed from the results of the survey from Phase 1. These workshops were designed to address salient concerns from the surveys, but

also to enable interactions and discussions that would help gain insight into the work practices of the participants as related to their use of language test scores.

## Participants

Workshops were held in eight institutions in six different provinces representing the Atlantic region, Central Canada, and Western Canada (Canada has 10 provinces and three northern territories). These eight institutions were evenly split between smaller primarily undergraduate institutions (with less than 10,000 undergraduate students) and larger more research-intensive universities (greater than 20,000 undergraduate students). The participants for the research study were 59 workshop attendees. Before the workshop, the facilitators and participants discussed ethical implications of the research, and participants provided signed ethical consent for their participation. It was made clear that participation in the research study was not necessary to attend the workshops. However, all but one workshop attendee consented to participate. The universities were asked to invite all those who worked directly with language assessment in decision making, in whatever role. As a result, participants came from a number of different units and job classifications. For example, sometimes English for academic purposes or language support personnel attended the workshops and sometimes they did not. It was very clear that some institutions have a close relationship established between international recruitment, admissions, and English language support, and some do not. Figure 2 summarizes the role distribution of the participants.

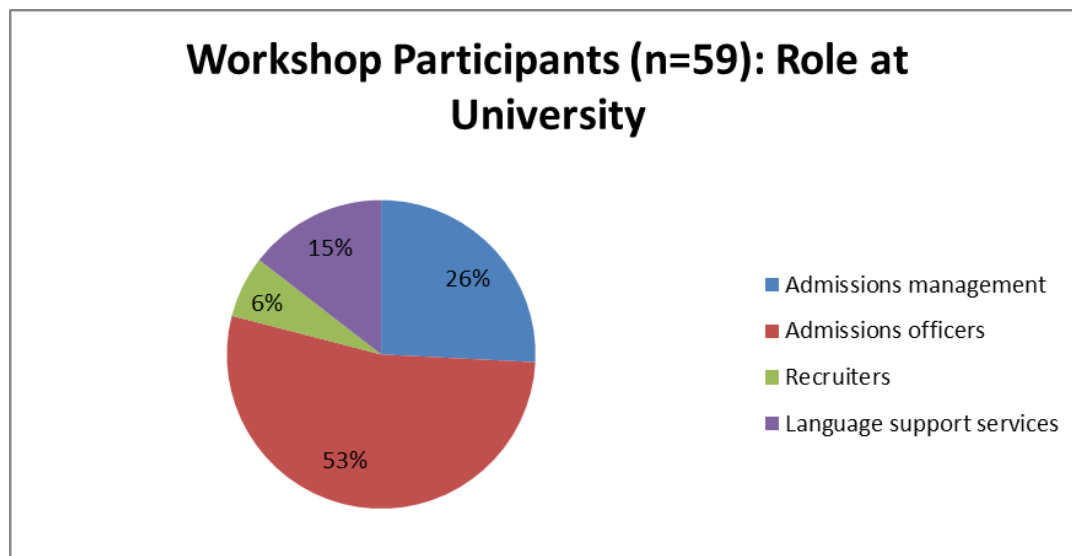


Figure 2. The roles of the workshop participants.

## Data Collection Procedures

The primary collection of data occurred during the workshop delivery, so it is important to describe the workshops themselves. Hands-on activities during these 2 ½ hour workshops included the following:

1. comparing most commonly-used exams on a number of factors—such as the cost of each test, the skills being tested, and the construct being addressed as stated by each test provider;
2. examining tasks and test taker samples, as well as grader comments, at levels that represented scores above and below their institution's current cutoffs;
3. comparing cutoff scores at various institutions and reviewing guidance from test providers on cutoff scores; and
4. discussing experiences with fraud and suspected fraud, reviewing related research and test provider information.

A full lunch and a coffee break were provided for the participants during the session, as both an incentive and a way of creating a more informal and conversational climate. During the workshops, our team (consisting of me and one or more research assistants) discussed characteristics of the most commonly used tests in nontechnical language. We also discussed non-testing pathways to admission, such as English as a second language (ESL) bridging programs. It is important to note that as workshop facilitators, we took the role of independently funded language test researchers who were not there to sell or otherwise promote the use of any one measure. We took pains to establish that all the major tests followed best practice in terms of their development and validation, without privileging one measure over another.

In these workshops, we also included information that explicitly countered any misconceptions we observed in the surveys, as well as common publically held notions regarding the infallibility of test scores (see Taylor, 2013). An example of this is Figure 3, which was a graphic used to illustrate the concept that test scores provide only some information on language ability in academic contexts, which in turn provides some information on their construct of interest—academic success. With this graphic, we were able to discuss issues of construct representation as well as construct relevance in jargon-free language. In our discussions, we appealed to our audience as professional decision makers, who had ostensibly learned through experience about the limitations of single measures in their work as well as the importance of having realistic expectations of these measures.



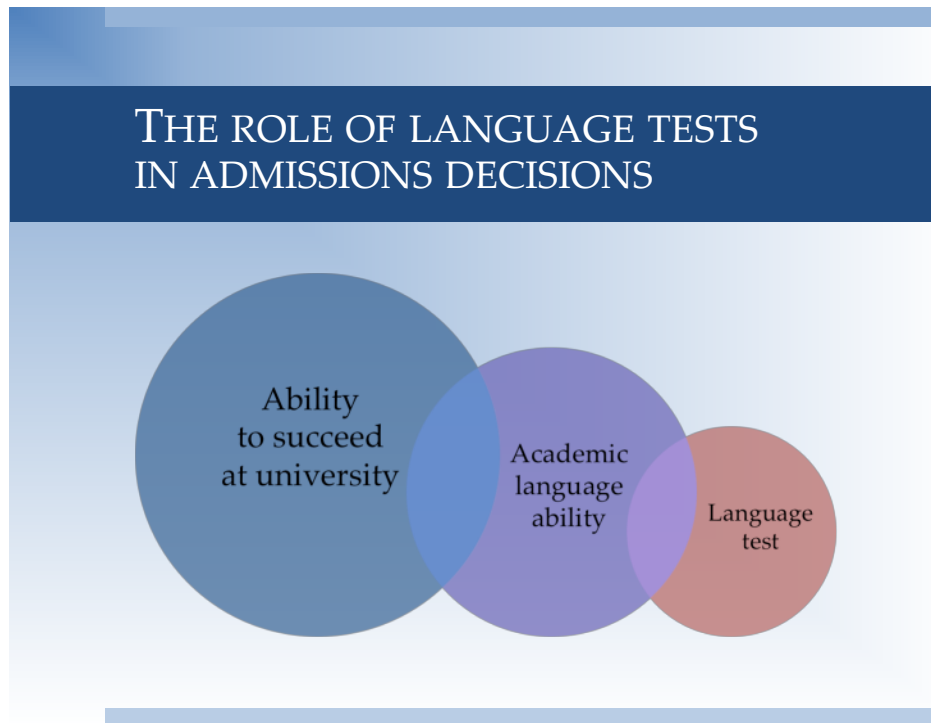


Figure 3. An example of a supporting graphic used during the workshops.

Data were collected during workshop sessions between May and October 2013. There were two data sources, each of which is explained below.

**1. Recordings of key “check-back” sessions during the workshop.** At three points during the workshop sessions, participants were asked to reflect on the previous activity and comment on what they learned or knew already, what surprised them, what they wanted to learn more about, or what they found the most pertinent to their work responsibilities.

Two research assistants from the research team were present for all workshops but two, and took notes during these recorded portions. One would then transcribe the session, and it would be verified and/or supplemented by the other. In two cases, only one research assistant was present so I performed the second reading and verification of the transcripts.

Participants were not asked what they learned about the field of language assessment, but what they learned which would be useful for their workplace performance. These discussions of what they learned, and how they would adjust their practices as a result, provided data to address the first research question by capturing information related to Eraut’s (2004a) typology. It was also predicted that some of these data would also address the second research question regarding the usefulness of this information in developing their professional competence.

**2. Responses and comments on a final workshop evaluation questionnaire.** Anonymous evaluation questionnaires followed each workshop. Besides querying the strengths and weaknesses of the workshop, the questionnaires included three 5-point Likert-type items which addressed our second research question, pertaining to the participants’

perceptions of skills and knowledge they may have acquired or enhanced during the session. In these questions, the participants indicated their level of agreement with the following statements:

1. I feel that I acquired new skills and knowledge.
2. I enhanced the skills and knowledge I already have.
3. The information presented will be valuable to me considering my current and/or anticipated job responsibilities.

Each of these questions had a space to provide additional comments, and there was a final space at the end of the survey for final comments. These comments were also compiled and combined with the transcripts from the workshop check-backs for analysis, with a view to providing additional evidence to address the second question primarily, but also perhaps this first question if tacit elements of Eraut's (2004a) typology were revealed indirectly.

### **Data Analysis Procedures**

Open-ended data were thematically analysed using NVivo 10 in an iterative process (Birks & Mills, 2011; Dörnyei 2007) in an attempt to apply the comments to the eight elements of the learning typology of Eraut (2004a). The transcripts were read while listening to the recordings during the first round of analysis, so that elements of the recordings not captured in the transcripts could be noted (such as a sarcastic tone).

During the coding process, it quickly became apparent that comments applied to multiple elements of the typology. It was decided to assign these comments to all elements of the typology to which they conceivably applied. For example, if a participant made a comment about wanting a summary of language testing research from us, this was viewed as evidence of an attention to research-based practice (i.e., academic knowledge and skills) as well as a request for outside expertise (decision making and problem solving). Following this process, a round of subcategorization was undertaken where each of the comments was analyzed for whether propositional or procedural knowledge was referenced, or whether the comments revealed explicit or tacit knowledge.

### **Results**

From the two data sources, there was evidence found for all elements of Eraut's (2004a) typology except for task performance. This was not surprising, as the data collected did not allow for the capture of real-time decision-making or other workplace tasks. The combination of small group sessions, a dedicated time away from work, as well as multiple reflective activities built into the workshops, appeared to have created a situation that allowed for rich data collection—in addition to the overall positive and open disposition of the participants.

As previously discussed, comments were often assigned to more than one element of the typology, but not in random ways—several pairs of categories often appeared together. Therefore, results of the analysis that address the first research question (i.e., an elaboration of the LAL of this group) are organized by the element or (much more often) the pair or group of elements that appear together. There will also be an identification of the

type of knowledge represented in the data (explicit or tacit, and procedural or propositional). Additional results are presented which address the second research question, that is, that speak to the participants' perceptions of the usefulness of the workshops to their professional development. All participant names are pseudonyms.

### **Awareness and Understanding**

Comments assigned to this category included empathetic comments about the student experience, in addition to awareness of institutional constraints and the balancing of multiple stakeholder interests. The knowledge in this classification is considered to be propositional knowledge. Workshop participants demonstrated empathy with students and awareness of ethical issues arising from their work, illustrated by the exchange below at the Institution "Small Central 3" regarding the burden of high costs of some standardized language tests in some areas of the world:

Joe (International Recruiter): [Standardized examinations are] pretty expensive.

Rita (International Admissions): It is expensive. ...

Mike (International Recruiter): It tells me if somebody puts the money up to take the test, they're serious.

Joe: Yeah, for us, we're like, "Oh, it's expensive." Imagine ... some countries it's a lot of money.

Participants often made comments where they put themselves in the position of the test taker, as the following example regarding the difference between the IELTS and the Test of English as a Foreign Language (TOEFL) speaking components:

Joan (Executive Director of Graduate Enrolment, "Large Western 2"): That's so fascinating. ... I am very extroverted but I probably might be more comfortable talking to a computer because it's not going to assess me.

Admissions decision makers also acknowledged local constraints regarding the setting of cutoff scores, related primarily to institutional pressures to increase student enrolment:

Joan (Associate Registrar, Admissions, "Large Central 1"): As an institution, when, if you have this score of 6.5—and you're hearing about all this data and research—like how would you actually [propose a new cutoff score] of 7 when, at the end of the day ... you need international numbers. ... You need to let them in the door ... So, um, raising them too high affects admission ... and when it implicates enrolment as a priority, that really has to be a balanced decision.

Charlie (Director, International Programs, "Small Central 2"): The pressure to provide that deliverable [increased international enrolment] is huge.

### **Role Performance, Decision Making/Problem Solving, and Judgment**

Comments were classified as relating to role performance if participants focused on their job requirements and responsibilities, and include discussions of procedures by which they handle ethical issues as part of their job. These comments are classified as representing procedural rather than propositional knowledge. These types of comments can also be interpreted as related indirectly to the use of information in decision making, as well as judgment (in that the consequences of decisions are discussed). This is illustrated in the following exchanges between Diane and Danielle at “Small Central 3,” and Yang and Suzanne at “Large Western 3,” where a problem is identified and the desire for additional information is expressed to allow for adjustments in current decision-making procedures:

Diane: (Admissions Officer): And if it [doesn't] work then, you know, there are students who are really coming in with lower scores and they're still not doing well ... then we really need to know. Why ... what aren't they doing? What aren't we doing to help them?

Danielle (Admissions Officer): Or if we don't have the resources to help them ...

Diane: Exactly.

Danielle: Then we need to be adjusting our initial [decision].

Yang (Admissions Director): What kind of support an institution can provide to students here? We are taking a risk as students are taking a risk. They come and if the student indeed cannot perform at the exact level, do we have the ESL program to send them to, right?

Suzanne (International Admissions): That is our job here. We're mandated to ensure your success; therefore, we need to give you what you need to be successful.

### **Academic Knowledge and Skills**

In the workshops, the admissions decision makers discussed their propositional knowledge of assessment concepts, such as the growing understanding of the complexity of the construct that the tests are attempting to capture, as in the following exchange at “Large Western 2”:

Janet (Academic Director): Academic English is not the same, I mean you don't just speak academic English like air traffic controllers only speak air traffic English.

Allison (Admissions): I'm trying to think, “So what does it mean to speak academic English?”

Janet: So that's a very odd term. Yes, I suppose they're trying to articulate some kind of complexity beyond conversational English, but ...

In addition, in several workshops on separate occasions, participants demonstrated their awareness of the limitations of a single test score in providing information about a complex concept (university success).

Max (Director, Admissions, “Large Western 1”): But the test score in itself shouldn’t be relied on all too heavily. It can’t determine the overall success in university. It’s, as you say, a number of tools.

Kimberley (International Admissions, “Small Eastern 1”): I take away from the workshop that [a language test] can give you an idea of what the language ability is like, but that doesn’t necessarily mean that they can do well at university.

Danielle (Admissions Officer, “Small Central 3”): The power of any one piece in an admissions file should not be overextended. It’s just one more ... it’s a piece.

Some comments reveal not understandings but common misconceptions, such as an underestimation of the amount of time needed to improve test scores:

Frederica (Director, Admissions, “Large Western 1”): How much students can evolve in their language ability [in a month]? .5 of a band [on IELTS], OK, it’s reasonable to assume that maybe in a month they could theoretically move into that band.

In the following comment Mike (International Recruitment, “Small Central 3”) reveals his previous tacit perception of assessment as a wholly objective process:

Mike: [Language assessment] seems very subjective—more subjective than I expected it to be.

Another illustration of previous tacit knowledge being made explicit is below, where Rita (Admissions Officer, “Large Central 1”) discusses her realization that her own intuitive estimations of an applicant’s language ability based on personal interaction may be potentially misleading. This comment reveals the tacit knowledge she had possessed previously—that initial intuitive judgments are reliable for decision making.

Rita: If you’ve met with somebody in person, you’ve had a conversation, [before] I would be willing to say, “Well, I think, in talking and understanding what he’s saying, I think he’s going to be okay.” But it’s a much more complex assessment than that.

### **Academic Knowledge and Decision Making/Problem Solving**

In discussing what they have learned in the workshop, participants pointed to propositional knowledge regarding the complexity of the construct being measured, in addition to discussing how their future decision-making behaviours would change as a result of this knowledge. These comments therefore link conceptual knowledge in assessment to the work activity, which involves synthesizing multiple sources of information in decision making. Therefore, both propositional and procedural knowledge are represented.

Paula (Language Support, “Large Central 1”): Well of all things [language ability] is one of many factors, right? And I think that’s, this is actually a big help for us with admissions as well just cause right now we are [considering] a lot of students on many different factors. And English language capabilities is one of them. I think that’s the relationship. ... It’s one aspect that we have in our toolbox of different things that we use to assess students’ success at the university.

### **Teamwork and Decision Making/Problem Solving**

Comments categorized as teamwork did not appear on their own, but always within comments related to procedural decision making. “We” is used almost exclusively in discussions of policy making and admissions decision-making procedures, suggesting that these decision-making processes are highly collaborative.

Manisha (Admissions Advisor, “Large Western 1”): The question is if we were to see someone submit two exams in a short period of time, and we see a huge variance in score. Should we naturally be suspicious of that?

Suzanne (International Admissions, “Small Eastern 2”): [S]o when we set a guideline for our cutoffs and are trying to predict success we have to think about what kind of ESL support there is. We also have to think about other things we don’t control, we can’t control, and that’s the environment they’re going to live in, the financial stress they may be under, and all those things.

### **Decision Making/Problem Solving and Judgment**

Participants discussed specific decision-making scenarios that revealed tacit procedural knowledge:

Irene (Director of Graduate Admissions, “Large Western 2”): Registrars deal with this issue all the time because, you know, I can give you a great example of getting a request from a student who had 1.98 and they needed 2 ... So you’ll get these requests, “Can I graduate?” And so it’s a very difficult kind of thing and yeah, anytime you have cutoff in assessment issues it’s very challenging.

Participants commented on their decision making with a consideration of the impact of their decisions, or the consequences of getting their decisions wrong. This is evidenced in the following exchange at “Large Western 3,” after the project’s research assistant mentions to the group that as an international applicant she was admitted to a Canadian university despite a very low speaking score on her TOEFL:

Bill (Admissions Officer): OK. So, it means that you’d rather not accept someone who might have been OK than accept someone who ends up not being OK.

Suzanne (Admissions Officer): I think so ...

Lauren (Admissions Officer): [But] I am thinking in the case of [the research assistant] ... we would not have taken *her*.

## Personal Development

Comments related to personal development primarily included the disposition to consult relevant knowledge and to continue learning, as is revealed by the following anonymous comments from the workshop questionnaires:

- “I’m interested in academic success & appreciate tests are not a predictor, but do scores align with our English testing—for me to research.”
- “The information presented has motivated me to do a number of studies related to our own students’ academic backgrounds, English assessment tests, & present academic performance.”

## Participants’ Perceptions of the Usefulness of the Workshop

The workshops were very positively received. The surveys performed in the prior phase of the project were worthwhile in that they enabled the creation of workshops that targeted the key preoccupations of the participants. Table 1 reports on the raw responses to the questions on the final anonymous questionnaire related to the perceived usefulness of the workshop.

Table 1

*Questionnaire Responses on the Usefulness of the Workshop*

Question	Very much so	For the most part	Somewhat	Only slightly	Not at all
The information presented will be valuable to me considering my current and/or anticipated job responsibilities.	.50	.32	.15	.03	.00
I feel that I acquired new skills and knowledge.	.29	.44	.24	.03	.00
I enhanced the skills and knowledge I already have	.29	.46	.24	.01	.00

## Discussion

### Research Question 1: What is the LAL needed for users of language test scores in admissions decision making at postsecondary institutions in Canada?

I have attempted to address this question from the point of view of professional competency development, including the extent to which the data represent propositional or procedural knowledge. Propositional knowledge in these data was related to the following:

- awareness of institutional constraints and the local context, including the various stakeholders involved and their (sometimes contradictory) interests;
- awareness of the consequences of participants' decisions on stakeholders;
- knowledge of concepts related to assessment more generally, such as the limitations of single test scores in capturing sufficient information about a complex construct, and the importance of rater judgment in performance assessment; and
- knowledge of concepts related to language assessment more specifically, such as the different conceptualizations of language proficiency for academic and general purposes.

Procedural knowledge as demonstrated in the data was primarily related to processes of judgment and decision making:

- simultaneous consideration of multiple alternatives, predicting the consequences of each;
- the use of heuristics (Gigerenzer & Gaissmaier, 2011) such as anchoring and adjustment (e.g., starting with a preliminary intuitive decision on language abilities based in personal meetings and adjusting based on ongoing information collecting);
- risk assessment; and
- determination of the point at which additional information must or should be sought to proceed.

A thorough treatment of the large and diverse literature in judgment and decision making (JDM) is beyond the scope of this report. However, JDM is being increasingly applied to domains such as institutional decision making (see Connolly, Arkes, & Hammond, 2000; Koehler & Harvey, 2004), and its usefulness suggests itself strongly here. If LAL competency for this stakeholder group includes procedural knowledge related to JDM, then a better understanding of JDM should be undertaken by all parties, in order to establish how the decision-making process can be improved upon (see Dhami, Schlottmann, & Waldmann, 2012).

Taylor's (2013) AL/LAL differential model, and in particular the profile for university administrators, is taken up here again in light of the results of this study. Taylor (2013) has questioned whether the construct of LAL should be considered as knowledge, a skill set, a set of principles, or all three. It seems to be a productive enterprise to operationalize LAL for this group as *competence* in the professional practice of using language test scores effectively in admissions decision making, and competence as defined here has included both propositional and procedural knowledge. Taylor (2013) has recommended further work to examine the usefulness of this model under empirical conditions, and this is an attempt to do so—to consider to what extent Taylor's (2013) categories for this group were represented in the data. This is summarized in Table 2.



Table 2

*Elements of Taylor's (2013) Differential Model for University Administrators, as Represented in the Data of This Study*

Elements of Taylor's (2013) AL/LAL Differential Model for University Administrators	Results of This Study as They Align to These Elements
Knowledge of theory; principles and concepts	propositional knowledge (both explicit and tacit) related to assessment in general and language assessment specifically
Technical skills; scores and decision-making	procedural role performance, both individually and in teams; skills in collaborative judgment and decision making (both explicit and tacit)
Language pedagogy	for these stakeholders, only propositional academic knowledge (such as the appropriacy of certain courses for applicants)
Local practices	awareness and understanding: propositional knowledge of the local situation and context
Personal beliefs and attitudes	propositional knowledge: underlying personal beliefs about their role (such as the obligation to serve students) and necessity for personal development

The last element of Taylor's (2013) model, sociocultural values, is not presented above and might instead be considered not as a separate element but permeating all other elements: Sociocultural values are often tacit and embedded in both our conceptual understandings as well as in our actions, such as our decision making. Figure 4 represents the resulting adapted profile for this stakeholder group. It is hoped that future work in LAL will examine to what extent this formulation (or one like it) can inform competency development.

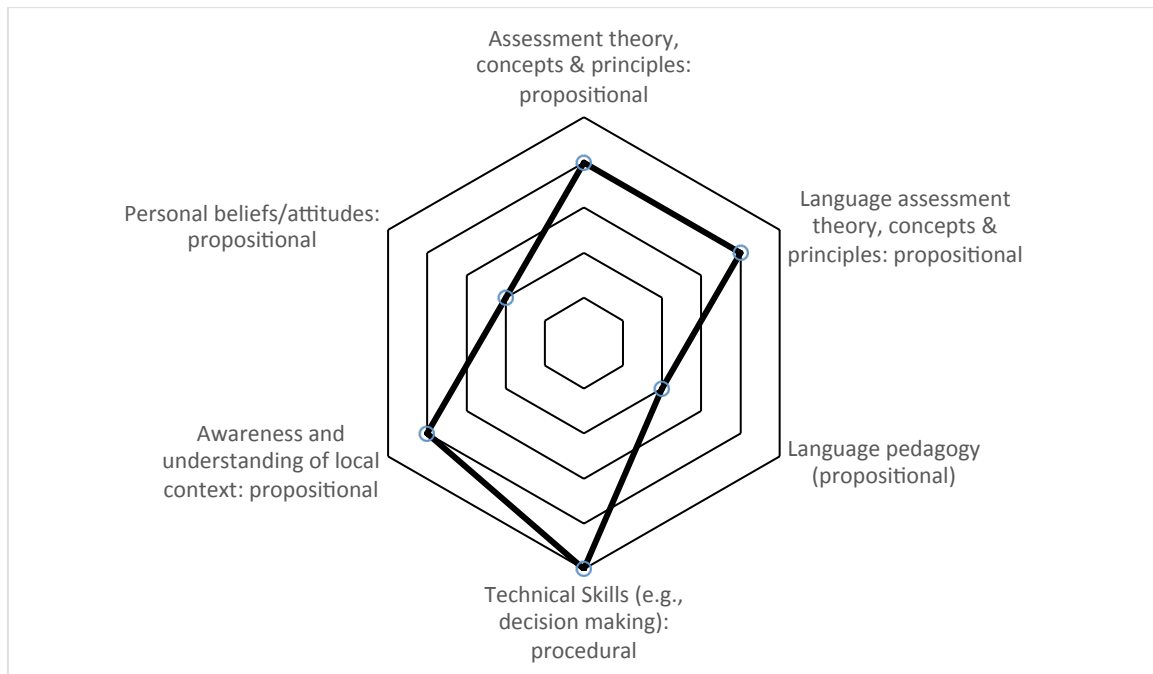


Figure 4. An adapted AL/LAL profile for university administrators.

Taylor's (2013) work includes the terms *assessment literacy* as well as *language assessment literacy* in describing this profile, and these terms are retained here: These data suggest that some theoretical and conceptual knowledge is related to assessment more generally, and some is related to language assessment concepts in particular.

Eraut and his colleagues (Eraut 1994, 2000, 2004a, 2004b; McKee & Eraut, 2012) have acknowledged that while many elements of their typology have tacit components, tacit knowledge is by definition not readily accessible, and may not even be recognized as part of workplace competence by workers themselves. However, in comments related to decisions and actions in their workplace, participants revealed tacit understandings (and misunderstandings). In fact, as the workshop was a learning experience, there was a rare opportunity to make explicit certain concepts. In those moments, participants revealed previous tacit misconceptions as they became clarified within the learning situation. It may be useful in future studies to attempt to capture tacit knowledge more fully with activities that allow for explicit discussions of the more tacit elements of their work.

### **Research Question 2: What useful materials can be created to develop this LAL for admissions decision makers?**

This question has only been partially addressed. Participants perceived the workshops and related materials to be extremely relevant to their practice. While these are important data, useful materials have to be not only be perceived as useful by all parties, but they must also lead to increased competence in the activity for which the scores are used. Future work must examine ways to empirically judge increased competence—how do we establish if LAL has improved/increased in addition to the decision makers' own conviction that it has? To assess development, performance information must be collected

at multiple intervals in a manner that takes into account the context and features of the performance, as well as indicators of expertise for the given performance.

While this work provides some evidence for the elaboration of an LAL base for these score users, collection of data during workshops meant collecting a snapshot of a dynamic process, not a static state. Knowledge, if we view it as imparted by the facilitators or arrived at by the participants, will be transformed as it is used, within cycles of use and reflection. Therefore, there is a limit to the usefulness of evaluating the benefits of a professional development activity during the event itself. This is not to say that perceptions of relevance are important to capture at the time or training, as well as evidence of reflection upon current attitudes, knowledge and practices. A good follow-up study would be to contact these admissions officers 2 years later and ask them to reflect on their interactions with language test scores in their work to sketch out a portrait of evolution—or lack thereof. However, a more complete evaluation of the development of a complex construct such as LAL would require the collection of evidence from multiple sources over time, and would need to capture data on real-time task performance.

Finally, even if the most useful materials for developing LAL are identified, the transfer of knowledge from education to workplace settings, and from previous experience to new situations, is much more complex than commonly perceived (Eraut, 2004a). Research in professional learning has suggested that the most effective methods involve a variety of settings, sufficient time for active and explicit reflection, accessible resources, and an openness to personal development.

### Conclusions

Previous work in LAL from admissions officers (Ginther & Elder 2013; Hyatt & Brooks, 2009; Rea-Dickins et al., 2007) has until now focused mostly on propositional or conceptual knowledge as defined by the language assessment community. This study provides some evidence that a competency-based conceptualization of LAL has the potential to add substantially to this previous work, by widening the conceptualization of LAL to include propositional versus procedural knowledge, as well as explicit and tacit elements. If LAL development is essentially professional competency development for these stakeholders, then it follows that improved LAL can also be assessed as improved job performance.

There is an additional benefit associated with expanding our conceptualization of LAL beyond formal propositional knowledge, one associated with collaborative knowledge sharing among stakeholders. If we as language assessment specialists concentrate on propositional knowledge of our field, we are always the ones with a superior position. Expanding the typology would include:

- elements language assessment specialists should take responsibility for sharing with other stakeholders, as well as
- elements about the contexts of use of language test scores that language assessment specialists should take responsibility for learning from other stakeholders.

In assessing improved job performance, the limitations of a language testing perspective must be acknowledged. Language testers produce a score intended for decision making, but we are not experts of decision-making processes and have a lot to learn from this literature

as well as from decision makers themselves. Other stakeholders must be therefore be included in delineating the construct of LAL.

Knowledge cannot be defined simply by the research community in the field, which leads too easily to the imposition of a deficiency model on non-researchers. Taylor (2013) also cautioned about the need to avoid arrogance in pursuing work in LAL. Eraut (1994) discussed how, in efforts to raise the reputation of their professional body, members of that body will tend to characterize their knowledge base as highly specialized and “carrying the aura of certainty” (p. 14). User-driven concerns can seem threatening, challenging already-established agendas:

A much broader framework is needed for studying the creation of professional knowledge; and the situation looks very different if we move the academic researcher from the centre of the universe. First we notice that new knowledge is created also by professionals in practice, though this is often of a different kind from that created by researchers. (Eraut, 1994, p. 54)

This research was conducted in this spirit.

Correspondence should be addressed to Beverly Baker.  
Email: Beverly.Baker@uottawa.ca

### **Acknowledgements**

I would like to thank the admissions officers and other university personnel for their valuable contributions to this study, as well as the members of the LAL Research Group (Rika Tsushima, Shujiao Wang, Mariusz Galczynski, and Sarah Desroches). This project was made possible through the Social Sciences and Humanities Research Council of Canada as well as McGill University.

### **References**

- Baker, B., Tsushima, R., & Wang, S. (2014). Investigating language assessment literacy: Collaboration between assessment specialists and Canadian university admissions officers [Special issue: Language testing and assessment in higher education: Some current issues]. *Language Learning in Higher Education*, 4(1), 137-157.
- Birks, M., & Mills, J. (2011). *Grounded theory: A practical guide*. London, United Kingdom: Sage.
- Canadian Bureau for International Education. (2014). *Canada's performance in international education, 2014*. Retrieved from <http://www.cbie.ca/>
- Connolly, T., Arkes, H. R., & Hammond, K. R. (2000). *Judgment and decision making: An interdisciplinary reader*. New York, NY: Cambridge University Press.
- Davies, A. (2008). Textbook trends in teaching language testing. *Language testing*, 25(3), 327-347.
- Dhami, M., Schlottmann, A., & Waldmann, M. (2012). *Judgment and decision making as a skill: Learning, development and evolution*. New York, NY: Cambridge University Press.

- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies*. Oxford, United Kingdom: Oxford University Press.
- Eraut, M. (1994). *Developing professional knowledge and competence*. London, United Kingdom: Falmer Press.
- Eraut, M. (2000). Non-formal learning and tacit knowledge in professional work. *British Journal of Educational Psychology*, 70(1), 113-130.
- Eraut, M. (2004a). Informal learning in the workplace. *Studies in Continuing Education*, 26(2), 247-274.
- Eraut, M. (2004b). Transfer of knowledge between education and workplace settings. In H. Rainbird, A. Fuller, & H. Munro (Eds.), *Workplace learning in context* (pp. 201-221). London, United Kingdom: Routledge.
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113-132.
- Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, 62, 451-482.
- Ginther, A., & Elder, C. (2013). *How are English language proficiency test scores understood and used by academic staff in the graduate admissions process?* Retrieved from [http://www.ealta.eu.org/conference/2013/presentations/Ginther%20and%20Elder%20presentat\\_on%20EALTA%202013.pdf](http://www.ealta.eu.org/conference/2013/presentations/Ginther%20and%20Elder%20presentat_on%20EALTA%202013.pdf)
- Hyatt, D., & Brooks, G. (2009). Investigating stakeholders' perceptions of IELTS as an entry requirement for higher education in the UK. *IELTS Research Reports*, 10, 17-68.
- Koehler, D. J., & Harvey, N. (Eds.). (2004). *Blackwell handbook of judgment and decision making* (Handbooks of Experimental Psychology). Malden, MA: Blackwell.
- McKee, A., & Eraut, M. (2012). Introduction. In A. McKee & M. Eraut (Eds.), *Learning trajectories, innovation and identity for professional development* (pp. 1-19). New York, NY: Springer.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (pp. 13-103). New York, NY: Macmillan.
- O'Loughlin, K. (2011). The interpretation and use of proficiency test scores in university selection: How valid and ethical are they? *Language Assessment Quarterly*, 8(2), 146-160.
- O'Loughlin, K. (2013). Developing the assessment literacy of university proficiency test users. *Language Testing*, 30(3), 363-380.
- Sharu, D. (2012). From nurse to advanced nurse practitioner: Mid-career transitions. In A. McKee & M. Eraut (Eds.), *Learning trajectories, innovation and identity for professional development* (pp. 175-192). New York, NY: Springer.
- Rea-Dickins, P., Kiely, R., & Yu, G. (2007). Student identity, learning and progression (SILP): The affective and academic impact of IELTS on "successful candidates." *IELTS Research Reports*, 7, 59-136.
- Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics*, 29, 21-36.
- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403-412.