



Contents

- Editors' Note [p. 1](#)
- Insights from Alister Cumming: An Interview [p. 2](#)
- A Highlight of LTRC 2018: Graduate Student Meet-up [p. 5](#)
- AALA 2018 Conference Coverage [p. 6](#)
 - Pre-conference workshops [p. 7](#)
 - Graduate student networking lunch [p. 8](#)
 - Words from student & dissertation award winners [p. 8](#)
- Dissertation Summaries [p. 10](#)
- 6th AALA Conference: Call for Abstracts [p. 12](#)
- AALA Information & Contact Details [p. 12](#)

Editors' Note

Ivy Chen & Kana Matsumura (Editors)

We're pleased to bring you the 4th issue of In Conversation (2018). As you may have noticed, this student newsletter, written by students for students, is now released annually rather than biennially and will cover the year's events as well as feature other engaging content from both fellow graduate students and well-established academics.

This issue begins with a follow-up interview with the keynote speaker (and workshop presenter) of last year's AALA conference, Alister Cumming, who shares his story and thoughts on the assessment of writing for academic purposes and gives us some great advice on professional development.

This year, the AALA student committee (SC) also collaborated with the student representatives of the Association for Language Testing and Assessment of Australia and New Zealand (ALTAANZ) to host an informal graduate student meeting at the Language Testing Research Colloquium (LTRC) held in New Zealand. One of the two SC members who attended reports.

Much of the newsletter pertains to our own AALA conference, which was held in Shanghai this year. We give an overview of the highlights, which is followed by short reports on two of the workshops our SC members attended (one on automated feedback and the other on rating scale development), and a summary of our graduate student networking lunch, where we were fortunate enough to have four academics come and share their experiences and advice for students. We also feature some words from the winners of the best student poster award, the best student paper award, and dissertation award.

Finally, we introduce a new addition to our newsletter content: dissertation summaries. We kick this off with a summary from our dissertation award winner but hope to have more of you share your research with us in future issues. (The SC will be sending out a call for summaries next year.)

We'd like to thank the SC for all their work this year and of course the wonderful organizers of this year's conference (Jin Yan and her team from Shanghai Jiao Tong University and those from FLTRP). Happy reading!

Insights from Alister Cumming: An Interview

Interviewed by Ivy Chen



Keynote speech at 4th AALA conference (Photo credit: LTTC)

Background: Could you share with us why and how you transitioned from a career in TESOL to a career as a researcher in the field of language testing?

Interesting question. I taught ESL to recent immigrants and visiting international students in Vancouver from the time that I completed a TESL Certificate in 1975 and then after moving to Montreal in 1979 (after completing a Master's degree in English) and then to Ottawa in 1981 to do similar work but focused on pre-university and first-year university ESL programs. While I was teaching at Carleton University in Ottawa three main factors influenced my decision to start a PhD in order to learn how to do research. First, after almost a decade learning how to teach English I felt relatively skilled at it and ready to develop new abilities and perspectives beyond the routines of classroom instruction. The second reason was that after precarious employment on yearly contracts, I realized that I would need to obtain a PhD to get a permanent position at any university in Canada. A key incident was my applying for a tenure-track position at Carleton University, for which I had all of the relevant teaching qualifications and experience except a PhD—and, of course, that position went to a person with a doctorate.

The third reason was that the department I worked in at Carleton University featured a number of active researchers who generated considerable excitement about their studies. I was inspired particularly by research that Ian Pringle, Aviva Freedman, and Stan Jones were doing on writing processes but also studies that Janice Yalden and my former colleague Ron Mackay from Concordia University were doing on curriculum development, needs analyses, and program evaluation. I came to realize how

little was actually known about language learning, writing, and education, and how important, useful, and fulfilling it was to conduct research that could provide principled, empirical guidance to practicing educators. But to pursue these interests I needed to take courses from skilled researchers and theorists and to get research experience myself, which is what I did from 1984 to 1989 as a PhD student at the Ontario Institute for Studies in Education. I have been intensively involved in research studies since that time. In fact, I am now finding it hard to stop as I move into a second year of retirement. Throughout my career, my primary interests have been on teaching and learning writing, but opportunities and needs have consistently arisen about assessing writing for academic purposes as well as in relation to immigrant settlement, I have found these activities (and the people and research associated with them) interesting and useful, and so I have followed them up.

“I came to realize how little was actually known about language learning, writing, and education, and how important, useful, and fulfilling it was to conduct research that could provide principled, empirical guidance to practicing educators.”

Research: A lot of your work has focused on the assessment of writing for academic purposes. Could you summarize for us how research in this area has evolved and what more can be done?

The assessment of writing for academic purposes has developed considerably over the past several decades, not only in terms of assessment practices but also in respect to research about them. At the very start of my career, the tests that language programs used at colleges and universities in Canada were simply of reading and listening comprehension, grammar, and vocabulary, and sometimes oral interviews. Tests like the CELT (Comprehensive English Language Test) and early versions of the MTB (Michigan Test Battery) were widely used to sort students into proficiency levels within language programs and to screen international or immigrant students for entry into university or college programs. In the late 1970s concerns were raised about the abilities of increasing numbers of international and immigrant students to write papers and exams effectively for university courses. In response, people and policies at universities started to implement direct tests of writing, that is, assessments that required prospective students to write compositions to demonstrate their abilities in English (rather than just to display their comprehension or knowledge of grammar and vocabulary). The purposes of these

assessments were primarily at a policy level for screening students into programs or placement into appropriate courses but also pedagogically for diagnosis of students' needs for learning or development within courses.

While at Concordia University I worked with Ron Mackay to develop a new test of writing in English for international students, for which Jack Upshur subsequently assumed responsibility and conducted research with Carolyn Turner, which is often cited for its binary-choice approach to raters' evaluation processes. Then while I was at Carleton University, Mari Wesche, Stan Jones, and others were developing the Ontario Test of English, which featured a direct test of writing in response to several reading passages, which was an early version of an integrated writing task and featured rating scales that were akin to the IELTS. These two exemplary projects coincided with influential work on the IELTS in the UK and former British colonies and in the USA by Holly Jacobs and colleagues on the ESL Composition Profile, Charles Stanfield on the TOEFL Test of Writing, and Liz Hamp-Lyons and Grant Henning on the Michigan English Language Assessment Battery. Together these projects instantiated several assessment practices as standard professional procedures for evaluating English writing: brief composition tasks that either involved writing arguments, describing graphical information, or responding to reading passages; descriptive criteria in the form of rating scales at a range of proficiency levels; and training raters to reach consensus on applying the rating scales to samples of compositions at relatively high levels of reliability.

As I mentioned earlier, there was a buzz about research on writing assessments in the late 1970s and early 1980s. Most of the initial research went into developing writing tasks, forming rating scales, and training raters. In subsequent years, I and many others realized that validation and refinement of these assessment practices needed to address additional considerations. One consideration was the decision-making processes that human raters applied to evaluate compositions. Although rating scales specify descriptive criteria about written texts at a range of proficiency levels, it is skilled human interpretations and judgements that actually evaluate and apply the criteria to make scoring decisions. Research involving think-aloud protocols, interviews, and questionnaires have proved useful to reveal and describe raters' thinking processes while scoring written compositions, served as a basis to enhance rater training, and can also explain otherwise intangible distinctions or trends in scoring.

A second consideration has been close analysis of written texts to identify and demarcate progressions or differences in language and literacy development. Various researchers have

analysed the morpho-syntactic features, range of lexical items, and discourse organization patterns in sets of compositions written at different levels of proficiency for language tests. The test validation issue here is whether the descriptive criteria in a rating scale account for systematic differences at each score point on the scale, whether additional linguistic criteria can be observed beyond those specified in the scoring rubrics, or even if construct-irrelevant variables (such as students' language backgrounds) appear. The wording of rating scales is necessarily broad and somewhat vague, however, and as observed above, their application depends on human impressions. So this approach to test verification is needed to ensure that the descriptors in rating scales are accurate and their interpretations during scoring are valid. Exciting prospects are also to be able to chart patterns of development for specific language features and their combinations in particular languages and tasks. Nonetheless, many studies have found great variability in the textual qualities of writing at specific score points, asserting the principle that proficient writing is multi-faceted, complex, and variable—in short, that good writing can be good in many different ways.

“[R]efinement of ... assessment practices needed to address ... the decision-making processes that human raters applied ... analysis of written texts to identify and demarcate progressions or differences in language and literacy development ... automated scoring ... ensuring that the writing tasks sampled in tests correspond to the types and qualities of writing that people need to perform”

A third, related consideration has been that scoring rubrics developed for human raters can be applied to computer programs for use as automated scoring. This research has involved the compilation and analysis of large corpora from assessment tasks to specify key lexical, morphological, syntactic, and discourse features that commonly feature at or distinguish language proficiency levels. Advances in this area are appearing regularly around the world, offering a valuable approach to test validation as well as substantial empirical insights about such things as the importance of collocations, formulaic phrases, and nominal phrases in academic writing. But as with all text analyses, such inquiry also reveals great differences and variability among individual writers as well as across rhetorical tasks and topics. Applications of sophisticated systems of automated scoring have proved to be as, or even more, reliable than human scoring, so can usefully complement or verify the work of human raters. Useful applications of automated scoring of

writing have been made for test practice, self-evaluation, and instructional purposes. A fundamental conceptual dilemma remains, however, that automated scoring asks people to write for machines rather than to communicate with other humans.

A fourth consideration was, and still is, ensuring that the writing tasks sampled in tests correspond to the types and qualities of writing that people need to perform for their academic studies or work. Initial studies of this kind involved various types of needs analysis, but their results have tended to be that almost everyone thinks that writing abilities are important and involve multiple dimensions of fluency, appropriateness, and accuracy. Recent studies have been comparing samples of written course papers and exams to the writing that students have done in tests, finding some common elements but also variability across contexts, topics, academic fields, and individuals. A glaring gap in current knowledge concerns describing and assessing the kinds of writing that people actually need to do for specialized types of work, for example, in the health professions, business or industry, or multi-media communications. Writing assessments have mostly been developed for large-scale contexts of academic study in secondary or higher education, aiming to generalize maximally from brief and highly conventional composition tasks, so formal assessments have rarely dealt with particular or unique genres of writing that are required for, and may be integral to, the performance of certain jobs.

Future directions for needed research and development are several. One is to enhance the diagnostic information from writing assessments for pedagogical purposes or self-study—while recognizing that diagnostic purposes are defined primarily in relation to particular educational programs and the specific opportunities for learning they are able to provide. A second area is to increase the range, genres, and samples of writing that examinees produce in tests to more fully represent an individual's abilities—despite the obvious constraints that writing is time consuming, people can only concentrate seriously for a few hours under testing conditions, and high levels of writing skill may not always be the most crucial ability that language learners want to develop. A third area, as I noted above, is the need to produce theoretically-substantiated accounts of writing development in second languages for particular learner groups and educational contexts. A related area that I think really warrants attention is to better define the construct of writing in a second language as distinct from writing abilities in a person's dominant language or mother tongue. To do so, cross-linguistic analyses are needed of people writing comparable tasks in their first and second (and additional) languages—a research design

that has been surprisingly absent to date in studies of second-language writing and assessment.

Advice: What advice would you give to our AALA student members with regards to their professional development?

Follow your interests because that is what will motivate, inspire, and define you as a professional. To do so, read, read, read, research, research, research, and write, write, write. Read as thoroughly, critically, and extensively as possible about the areas in which you decide to specialize. Although all aspects of language assessment have expanded considerably in recent decades, the extent of published writing about their research, theories, and key concepts remain definable and within the grasp of any person willing to spend time to master them. A broad perspective is valuable, but every young scholar needs to develop specialized expertise in two or three areas in order to be able to know what knowledge currently exists and needs to be researched further and to teach courses or develop assessments about them. The number of professional language assessors is relatively small, too, so opportunities readily exist to excel in and contribute usefully to specialized aspects of language assessment.

“Develop expertise with certain research approaches but also try out different, complementary methods of inquiry. Get skilled at research and feel confident about doing it.”

But to do so, one has to know the specialized area(s) thoroughly and in depth, and acquiring that knowledge requires reading comprehensively and purposefully. Particularly important to read are key

journals on language assessment and on educational assessment generally as well as key, commonly cited but also new, ground-breaking books and theories.

Get research experience while doing your graduate degrees by working with professors on their projects, conducting small empirical or case studies for course papers, and taking on part-time employment with testing or other assessment agencies. Develop expertise with certain research approaches but also try out different, complementary methods of inquiry. Get skilled at research and feel confident about doing it. Do your own analyses first hand, designing needed studies, gathering and analysing data, and interpreting and sharing the findings. Attend conferences like AALA, the Language Testing Research Colloquium, and larger venues such as the American Association for Applied Linguistics or American Educational Research Association. See there what experienced researchers are doing and saying

as well as what new innovations, concerns, and trends are emerging. Present your own research, as well, either as a work in progress or a poster or as a paper once you have completed a thesis or other significant project. In this way, develop and assert your professional identity, learn from others' questions and comments, and establish contacts with people who are interested in your work and ideas.

Above all, write about your research, ideas, and interests. If you have done a thesis for a Master's or Doctoral degree, persevere to get it published. A definition of a thesis is that it is a unique contribution to current knowledge worthy of publication. Going through the work to prepare a short manuscript from a thesis for a journal article can be arduous, time consuming, and even frustrating (and it still is for me, too). But publications are also rewarding, the necessary step toward establishing a career as a scholar, and the means of becoming a recognized member of the specialized discourse community called language assessors. I am always amazed how many people around the world read my

“Above all, write about your research, ideas, and interests. If you have done a thesis for a Master's or Doctoral degree, persevere to get it published.”

published research articles, cite them, and build on and extend their findings—even for decades afterwards. Use conference presentations, informal local colloquia, or study groups as media to try out your thoughts or results

of research initially, and then make a point of synthesizing those ideas into a form that is ready to be published, and be prepared to revise your work to satisfy the expectations and suggestions of reviewers and editors. Write also in other less demanding venues such as book reviews, newsletter articles, critiques of publications or theories, or even interviews (like this one) with established researchers whose work you admire. Graduate students who publish while and just after completing their doctoral degrees are invariably the ones who get the good jobs.

A Highlight of LTRC 2018: Graduate Student Meet-up

Report by Carol Dabarera

The 40th Language Testing Research Colloquium (LTRC) organized by the International Language Testing Association (ILTA) was held from 2 – 6 July 2018 in Auckland, New Zealand. The theme for the conference was “Language Assessment in Times of Movement, Transition, and Change”. For the very first time, the student committee (SC) of the Asian Association for Language Testing (AALA) together with the student representatives of the Association for Language Testing and Assessment of Australia and New Zealand (ALTAANZ) held an informal graduate student networking session at the LTRC. This

collaboration brought together two student groups with the common aim of reaching out to fellow students and to provide a platform for networking.

The session which was attended by both student presenters and participants of the LTRC began with an introduction of the associations and their student committees. Thereafter the attendees interacted with one another, where they very frequently connected over similar research strands and education experiences. The session ended well with the participants' initiation into a community of students enthusiastic about language research and in particular, language testing.



Student attendees at the student meeting at LTRC 2018 (Photo credit: Judit Tunde McPherson)

AALA SC representation: Carol, Xiaohua, and Ivy (3rd, 4th, and 5th from the left in the front row)

AALA 2018 Conference Coverage

Report by Ivy Chen

The 5th International Conference of the Asian Association for Language Assessment was held from October 18 to October 20 in Shanghai, China, hosted by Shanghai Jiao Tong University (SJTU) with associate host Foreign Language Teaching and Research Press. The theme for the conference was “Standards in Language Assessment” and the conference served as a tribute to Cyril Weir, a legend in language testing.

A big step up from last year, in large due to the location, an amazing 300 conference participants from over 19 countries and regions attended, among which there were 100 student participants from 10 countries and regions. The conference comprised of four plenary speeches in addition to the keynote speech, four pre-conference workshops, 54 papers, and 15 posters.

The pre-conference day consisted of four workshops: one on rating scale development with our keynote speaker, Ute Knoch (see p. 7 for a summary), one on automated feedback with Lin Gu and Keelan Evanini (see p. 7), one on using online resources

for identifying suitable reading texts with Barry O’Sullivan, and one on assessing reading with Hanan Khalifa. The day concluded with a cozy reception dinner.

The first day of the conference began with our keynote speech on the issue of context in scale design. This was also the day of our well-attended graduate student networking lunch (see p. 8 for details). We took a group photo at the end of the day (see below). Those of us who attended the banquet dinner enjoyed a wonderful cruise along the Huangpu River. On the second day of the conference, the poster session was held after lunch.

Lastly, we’d like to congratulate once more the winners of the best student poster award (Yangping Deng), the best student paper award (Kana Matsumura), and outstanding dissertation award (Yanfeng Yang). See page 8 for a few words from these winners and page 10 for Yanfeng Yang’s dissertation summary.

Overall, it was a well-organized conference with stimulating presentations.



Conference group photo (Photo credit: SJTU)

Pre-conference Workshops

Workshop 1: Automated Feedback on Spoken and Written Language Production

Report by Sha Liu

This report features “Workshop 1: Automated feedback on spoken and written language production” by Dr Lin Gu and Dr Keelan Evanini from Educational Testing Service (hereafter ETS) in the US.

The purpose of the one-day workshop was to enhance attendees’ knowledge of the development and evaluation of automated feedback on L2 spoken and written language production in language learning and assessment settings. As such, the workshop was designed to provide information in two main areas. The morning session focused on the use of automated feedback in learning contexts. In doing so, Dr Gu and Dr Evanini demonstrated a few research prototype feedback systems currently being developed by ETS. They also provided a critical review of the most common types of feedback generated by current automated feedback systems and previous research on the instructional application of such systems in learning contexts. The afternoon session addresses the validity of using automated feedback in assessment settings. Specifically, Dr Gu and Dr Evanini demonstrated how validity considerations have been addressed throughout the design of automated evaluation systems by ETS.

The workshop is a direct and timely response to EFL/ESL writing instructors’ and researchers’ urgent need for a better understanding of automated feedback systems. It was clearly-structured with a variety of engaging and well-thought-out activities. Throughout the workshop sessions, attendees were encouraged to discuss and share their understanding of the use of automated feedback systems for both learning and assessment purposes, which was thought-provoking and greatly enhanced the attendees’ understanding. Particularly, the provision of access code to Criterion, the automated feedback system developed by ETS, enabled the attendees to gain first-hand experience of using and evaluating such systems. Attendees were provided with sample essays in advance for the trial of Criterion during the workshop and with one-month access to the system afterwards. Such opportunity is particularly invaluable for writing instructors and researchers who desire to learn more about automated feedback systems but have difficulty in gaining the access.

Workshop 3: Rating Scale Development and Validation

Report by Tatsuro Tahara

This is the report of “Workshop 3: Rating scale development and validation” by Dr. Ute Knoch, Language Testing Research Centre, University of Melbourne.

The purpose of the workshop was to familiarize participants with rating scale development and validation. Participants were able to understand the differences among several types of rating scales as well as how to develop and apply the rating scales in their context. In particular, they learned how to integrate the rating scales into writing and speaking assessment into their teaching and learning. At the end of the workshop, participants examined methods to validate the use of rating scales by referring to Knoch & Chapelle’s (2018) latest framework of test validation.

Dr. Knoch facilitated the participants’ involvement in activities. First, she used various kinds of tasks so that participants could understand the uses of rating scales. For example, at the beginning of the workshop, Dr. Knoch showed the two types of rating scales (analytic scales & EBB scales) and asked participants the differences and merits/demerits of the use of rating scales. Participants paired up and discussed how to use them in their own classes, and later shared what they talked about with the floor. Moreover, Dr. Knoch asked some questions regarding how to validate rating scales, such as “How would you conduct validation of rating scales if there were unlimited resources to do researches?” Participants raised hands and shared their ideas in a lively manner. Through the activities, participants were getting to know how to develop or adapt the appropriate rating scales in their contexts as well as validation methods.

Graduate Student Networking Lunch

Report by Kana Matsumura



<Graduate Student Networking (GSN) session attendees>

The AALA Graduate Student Networking Session at the Annual Conference was inceptioned in 2016. It has been a platform for graduate students. The purposes of the GSN session is to network among ourselves and to learn from invited guests (i.e., EB members, professors, and professionals).

At this year's conference, there were 100 student participants from 10 countries and regions. (Isn't that great?) And about thirty students joined our session. We invited four professors to have them share their thoughts about research. We also had the participants talk in pairs or in groups to introduce themselves.

We had a raffle draw (books on applied linguistics contributed by invited guests/EB members) as usual as previous years.

The time was so limited, but we hope we had some time to get to know each other. We really appreciate your support! Let's keep in touch!



GSN lunch
19th, Oct. 2018

Words from Student & Dissertation Award Winners

Introduction by Ivy Chen

The two student awards were established in 2016 and are merit-based awards that support conference attendance. Our award winners were asked to write a little bit about who they are and their research interests and experiences. They then introduce the study that won them the award. I wanted to point out that the winner of this year's best paper award (aka my wonderful co-editor) also won the best poster award at last year's conference, so please refer to the last issue of our newsletter for more details.

We're also happy that we had our first ever outstanding dissertation award winner this year. If you're close to completion or know someone who has recently completed their dissertations (within three years), do apply for this award when they send out the call for nominations next year!

AALA Best Student Poster Award



I am Yanping Deng, studying in Waseda University for my PhD currently. My research interest is raters' bias in speaking test as there are many issues related to this topic. I completed my bachelor and master studies in Chongqing University. My previous study was the traditional culture reflected in Japanese classical literary works. Then I turned my focus on the research of learners' learning strategies and HSK (Chinese proficiency test) five years ago. Last year, I came to Waseda University to accept formal and systematic education of language testing.

My research study for the 5th AALA was to investigate Chinese students' use and interpretation of the TOEFL rubrics. The results showed that the three student-raters had three rating patterns which may result from their learning experience and test-

taking experience. They had similar rating patterns when rating themselves and others' speech samples. The results gave me a new hint that for nonnative English speaker raters, their testing-taking experience may affect their rating processes as well. This research for AALA is a novel idea, where the raters are the students who are preparing the TOEFL speaking test. They usually rate peers' work during preparation. They play two roles during the preparation: learners and raters. It is very different from the raters described in previous studies. Therefore, I could get reliable information from these students to consider the validity of the speaking rubrics.

I am a beginner in language testing. Moreover, it was the first time for me to participate in AALA. Therefore, I felt very surprised and honored to get the award as I never expected it. It gave me great motivation and confidence on my research. I really like to collect data and use quantitative method to analyze the data. I have failed a few times in collecting data, however, I could learn much from my failure experience. It is an exciting thing to have an idea and apply it in my research.

My main topic is to investigate the sources of raters' bias, especially raters' linguistic background and try to minimize the bias. If you have interest in speaking test, rater bias, G theory, and Rasch model, etc., please contact me. My email address is totodeng117@akane.waseda.jp.

AALA Best Student Paper Award



Hello AALA student members. I am Kana Matsumura, a graduate student in Waseda University, Tokyo, Japan. I am also a Student Committee member of AALA. I am very much honored to have won the 2018 Best Student Paper Award. I also won the Best Student Poster Award last year and I am very happy to have been involved in AALA in many ways!

My current research interests include EFL assessment of argumentative performances in the classroom context in particular. I am working on my Ph. D. dissertation to analyze Japanese EFL students' argumentative writing abilities based on Toulmin model of argument by applying mixed methods approach. I have been teaching English for 15 years in high schools and now I also teach at some universities. I have been always thinking of how teachers in classroom can provide our students with efficient and economical feedback on their performances.

The title of my presentation is "A Generalizability Analysis of Japanese EFL Students' Performance on Argumentative Writing Tasks". The present study attempts to examine the dependability of the rating scales using G-theory (Brennan, 1983; Cronbach, Gleser, Nanda, & Rajaratnam, 1972; Shavelson & Webb, 1991) which can investigate the relative effects of multiple sources of variance in the test scores such as tasks and rater judgments (e.g., Bachman, 2004; Cumming, 1990; McNamara, 1990). The rating scale on the part of argumentative content was developed by the author based on the framework of the Empirically derived, Binary-choice, Boundary-definition (EBB) approach (Koizumi & Hirai, 2008; 2013; Turner & Upshur, 1995; 1996; 2002). The language part was based on the Empirically-derived Descriptor-based Diagnostic (EDD) checklist (Kim, 2011) partially modified by the author.

Generalizability theory (G theory) is, as Lynch and McNamara (1998) note, "useful in providing general, group-level information, and particularly in making overall decisions about test design (p.176). This is why I am very much fascinated by the framework as an instructor and a test designer in classroom as well.

Again, I am very much honored to have won the award and I want to show my acknowledgement to my supervisor, Dr. Yasuyo Sawaki, who always gives me the best advice. I also want to thank AALA as well for being my research platform.

kmatsumura@suou.waseda.jp

AALA Outstanding Dissertation Award



Hello AALA student members! I am Yanfeng Yang, a graduate from The Hong Kong Polytechnic University. I am extremely honored to have won the 2018 Outstanding Dissertation Award of AALA and very happy to share my story with you.

My current research interests include sociocultural theory and its applications to L2 education, e.g., dynamic assessment (DA) and concept-based instruction (CBI). I have been teaching English for 10 years since I earned my MA degree. I am currently an associate professor and the team leader of applied linguistics in the School of Foreign Studies in Lingnan Normal University.

Dynamic assessment is a challenging topic for me at the beginning. I am greatly indebted to my supervisor Professor David Qian. It is he who introduced me to this interesting topic and provided me with enormous support throughout my research. His expertise, diligence, wisdom, and kindness impressed me a lot and motivated me to go forward. I am also grateful to the reviewers for their valuable and insightful comments. My gratitude also goes to my teachers, friends, colleagues, participants, and family. Without their support, I would not have completed my thesis. Thank you all!

Dissertation Summaries

As we explained at the beginning of this newsletter, this section is a new initiative designed for us to share our research with other students, so that we can learn and gain inspiration from each other. Although we only have one summary here this time, from Yanfeng Yang, the winner of the Outstanding Dissertation award featured above, we'd love to hear from everyone and include more of your research in our future issues.

Promoting L2 English Learners' Reading Abilities through Computerized Dynamic Assessment

Abstract by Yanfeng Yang

Derived from Vygotsky's proposition of the zone of proximal development (ZPD), dynamic assessment (DA) is distinguished from traditional testing in that the assessor intervenes in the assessment process to help learners perform beyond their independent functioning and concurrently promote learner development. DA not only reveals the learners' actual level of development through their independent problem solving but also their potential level of development through the interaction between the learner and the mediator, the latter of which is more useful for future learning and instruction. However, the number of learners a DA assessor can help is limited due to the time demand on mediator-learner interactions. To overcome this weakness, researchers start turning to computerized dynamic assessment (C-DA). Despite valuable insights gained from the few C-DA studies in second language education (e.g., Poehner & Lantolf, 2013; Poehner et al., 2014), the effectiveness of C-DA is still unclear. Additional studies are needed, especially longitudinal studies of between-group design. Reading comprehension is an important skill in language learning but it is also a hidden and complicated process which is difficult to

Last but not least, I would like to thank AALA for providing such a wonderful platform for researchers from all over the world to share their studies and ideas at the AALA conferences. I benefitted a lot from being an AALA student member and am very thrilled to eye-witness the rapid growth of AALA in the past few years. Thank you AALA! Wish you a more prosperous future! Finally, I would like to give a word to AALA student members and to myself — "Never give up pursuing your dreams and enjoy the happiness the research brings to you".

observe. Alderson et al. (2015) demonstrated the capacities of DA in diagnosing learners' strengths and weaknesses in second/foreign language reading. However, there is a paucity of related literature.

Targeting at Chinese EFL learners' reading comprehension, the present research used C-DA as a teaching and assessment method to promote learners' reading comprehension proficiency. Following a mixed methods design, the present project consisted of two studies. Study 1, a preliminary study, aimed to test the appropriateness of the C-DA program developed for this project. Study 2 consisted of two phases: Phase 1 and Phase 2. The main purpose of Phase 1 was to identify, through one-on-one interactions between the mediator and several learners, appropriate mediational strategies that could promote learner development. The findings from Phase 1 served as the basis for writing the C-DA mediation content in Phase 2. Informed by the results of Study 1 and Phase 1 of Study 2, Phase 2 of Study 2 investigated the extent to which C-DA promoted learner development as compared with the effects of traditional methods of teaching and assessment in reading comprehension.

Study 1 consisted of two reading tests, one in traditional multiple choice questions (MCQs) and the other in C-DA. The former was administered to a group of students who were asked to write down their thinking processes of doing the test, from which the students' difficulties in reading comprehension were identified and mediational strategies were worked out for the C-DA test. Another group of students completed the C-DA test, following which a questionnaire was distributed in order to obtain the students' reflections on the C-DA program. It was found that C-DA had advantages over traditional static assessment in diagnostic capacities and depth of insights an assessment procedure could provide for subsequent teaching. The results of the questionnaire showed that

C-DA was greatly welcomed by the students and believed to be particularly helpful in building up confidence for low achievers as identified in traditional static assessment.

Phase 1 of Study 2 was conducted among five participants and lasted for 7 weeks. The instrument consisted of MCQs selected from the reading section of the original TOEFL pBT tests. The MCQs were separated into three tests according to the item difficulty levels based on a series of tests in a pilot study: pre-test, post-test comparable to pre-test in terms of the item difficulty level, and transfer test which was more difficult than pre- and post- tests. The three tests were administered in 1st, 5th, and 7th week respectively. Following each test, one-on-one DA interactions were conducted between the mediator and the learner. The results of the pre-test and the problems identified during the DA interactions served as the baseline for the subsequent four-week enrichment program (EP). Comparisons of the students' independent and mediated performances on the pre- and post-tests showed whether or not the students benefited from DA and EP. Comparisons of the students' independent and mediated performances on the transfer test with those on the previous two tests indicated to what extent students can regulate their functioning when the context changed. Results showed that DA was effective in promoting learner development in their reading comprehension. Mediation moves deemed helpful for the DA process were also identified.

Based on the research findings of Phase 1 of Study 2, especially those from the DA interactions between the mediator and the

students, the mediations were written for each item of the three tests and applied to the large-scale study through the implementation of C-DA. In order to get additional evidence for the effects of C-DA, between-group design including an experimental group and a control group was adopted in Phase 2 of Study 2. Like in Phase 1 of Study 2, both groups were administered a pre-test, a post-test, and a transfer test, but the tests for the two groups were in different formats. The C-DA was used for the experimental group whereas traditional MCQs for the control group. The two groups also differed from each other in their four-week EP. The experimental group was instructed using the results of DA and through the employment of C-DA while the control group was taught using the traditional teaching method. The results of the three tests were compared within each group and between the two groups using repeated measures mixed factorial analysis of variance, from which the advantages of C-DA over the traditional teaching and assessment method were revealed. Results showed that the two groups were very close in their reading comprehension performances at the beginning of the study, but their improvements varied significantly after the four weeks of the intervention; the experimental group performed significantly better than the control group.

To sum up, the present research provides rich insights of the advantages of C-DA over traditional teaching and assessment methods in promoting second language learning, contributing to Vygotsky's social-cultural theory and ZPD, as well as knowledge of DA and C-DA in the field of L2 research.

Call for Abstracts: The 6th AALA International Conference

We are now welcoming abstracts for the 6th International Conference of the Asian Association for Language Assessment (AALA), Hanoi, October 17-18, 2019 and invite contributions from academics, researchers, teachers, students, and assessment agencies.

- Theme: **Language Assessment: Principles, Policies, and Practices**

- Conference Dates: October 16-18

- Pre-Conference Workshops: **October 16**

- Host: University of Languages and International Studies, Vietnam National University, Hanoi

- Associate hosts: University of Foreign Languages, Hue University, University of Foreign Language Studies, the University of Da Nang

- Co-Chairs: Jessica Wu & Jin Yan

- Local Chair: Nguyen Thi Ngoc Quynh

- **Deadline for submission: March 31, 2019**

- Notification of selection: June 30, 2019

For more information, see the conference website:

<http://ocs.ulis.vnu.edu.vn/index.php/aala/aala2019>.

AALA Founding Members (in alphabetical order)

He, Lianzhen (Zhejiang University, China)

Imao, Yasu (Osaka University, Japan)

Jin, Yan (Shanghai Jiao Tong University, China)

Kunnan, Antony John (University of Macau, Macau)

Lee, Yong-Won (Seoul National University, South Korea)

Lee, Young-Shik (Hannam University, South Korea)

Qian, David (Hong Kong Polytechnic University, Hong Kong)

Sawaki, Yasuyo (Waseda University, Japan)

Wu, Jessica Row-Whei (Language Training and Testing Center, Taiwan)

AALA Executive Board (2018-2019)

President

Jessica Wu (Language Training and Testing Center)

Co-President

Jin Yan (Shanghai Jiao Tong University)

1st Vice President

Jirada Wudthayagorn (Chulalongkorn University)

2nd Vice President

Nguyen Thi Ngoc Quynh (Vietnam National University)

Immediate Past President

David Qian (Hong Kong Polytechnic University)

Secretary/Treasurer

Yasuyo Sawaki (Waseda University)

Communications Officer

Yasuhiro Imao (Osaka University)

Members at Large

Yo In'nami (Chuo University)

Alan Urmston (Hong Kong Polytechnic University)

AALA Student Committee (2018-2019)

Advisor

Jirada Wudthayagorn (Chulalongkorn University, Thailand)

Co-chairs

Sha Liu (University of Bristol, UK)

Ivy Chen (University of Melbourne, Australia)

Immediate Past Co-chair

Rongchan Lin (Nanyang Technological University, Singapore)

Secretary/Treasurer

Carol Dabarera (Nanyang Technological University, Singapore)

Public Relations Officer

Kana Matsumura (Waseda University, Japan)

Committee Members

Tatsuro Tahara (Waseda University, Japan)

Committee Contact Information

AALA Website: <http://www.aalawebsite.com/>

Email: aalastudent@hotmail.com

Facebook Page:

<https://www.facebook.com/aalastudentcommittee/>

Facebook Group:

<https://www.facebook.com/groups/AALASTudentGroup/>

LinkedIn:

<https://www.linkedin.com/groups/6941774/profile>

AALA Student Committee *In Conversation Series*

Issue 4 (2018)

Editors: Ivy Chen & Kana Matsumura