AJAL

Preparing Chinese EFL adults for life abroad: Self-assessment, listening, and speaking

Haley Dolosic, Wei-Chieh Fang, Cindy Brantmeier, Michael Strube
Washington University in St. Louis, USA

Yanming Gao
Northeast Normal University, China

This study examines relationships among self-assessment and a high-stakes listening and speaking exam with 82 adult Chinese scholars seeking scholarships from the China Scholarship Council to support their research in the English-speaking academy. Students completed a criterion-referenced self-assessment questionnaire (Brantmeier, Vanderplank, & Strube, 2012) and a high-stakes test of English at the end of a rigorous English training programme aimed to prepare students for life within the English-speaking academy. Correlation analyses indicated correlations between listening and speaking performance ($r = .41$, $p < .01$) on the exam. Likewise, correlations between self-assessed listening and speaking were statistically significant ($r = .76$, $p < .01$). Yet, there is a statistically significant difference between the performance and self-assessed correlations, meaning students' self-assessments of listening and speaking are significantly more closely associated than their actual listening and speaking performance. Further, students' self-assessed listening and speaking abilities did not significantly correlate with performance measures on the high stakes test of listening and speaking ($p > .05$). Together, these findings suggest that these learners were not able to accurately depict their own strengths and weaknesses in listening and speaking. These results are examined with reference to the unique learning context, providing insights into how practitioners can encourage Chinese EFL learners to better self-assess.

Keywords: EFL; China; self-assessment; language learning; adult learners

Introduction

Within China the population of individuals studying English as a foreign language (EFL) is growing (Li & Cutting, 2011). This is partly due to English being the first foreign language taught in middle schools in China (Wu, 2016), but also because millions of other Chinese citizens including soldiers, business people, and government officials are endeavouring to learn English (Lai, 2001) to equip themselves with the skills needed to pass exams, thereby affording access to more desirable positions in society both within and outside China (Cheng, 2008). Yet, when these EFL learners seek to traverse linguistic and national boundaries for their own social, entrepreneurial, and academic pursuits, they may not be aware of their own capabilities to succeed at these language exams (Dolosic, 2018; Schultz, 2017). Such self-knowledge could be crucial for lifelong language learning (Brantmeier et al., 2012), because learners who are involved in the assessment of their own language skills are able to become more autonomous learners (Little, 2009; Nguyen & Gu, 2013). Prior research with Chinese EFL students has made substantial progress in understanding university classroom activities that promote listening and speaking proficiency (Gan, 2012) and the relationship among affective variables and the skills of

listening and speaking (M. Liu & Jackson, 2008). Yet, despite the benefits of self-assessment, few studies have examined whether learners who study EFL in a high-stakes academic training programme are able to diagnose their own strengths and weaknesses in EFL speaking and listening, particularly within mainland China. Therefore, the present study examines the relationships between adult Chinese learners' listening and speaking self-assessment and their performance following an intensive academic English language training programme. The goal is to understand whether these learners are able to depict their own abilities with accuracy.

Literature review

Listening and speaking
Listening and speaking are foundational to learners' successful participation in academic and professional settings, shaping individuals' abilities to both understand and communicate complex ideas in real-time interactions (Hughes & Reed, 2016; Rost, 2002). Speaking is a communicative, interactional competence by which learners can produce not only verbal output but deliver this output in a socially appropriate manner (Hughes & Reed, 2016). This complicated process happens quickly with the speaker deciding what to say and forming the words almost instantaneously (Hughes & Reed, 2016). Within many communication paradigms, this speaking skill is critically associated with the concept of listening. Listening is considered to be the act of attending to perceived sounds, segmenting and parsing words and phrases, and integrating the information these words and phrases carry into prior knowledge bases (Rost, 2002). This complex process relies on the listener attending to the verbal and physical cues of the listening experience, using their pragmatic and language knowledge to understand the intended meaning of the speaker (Rost, 2002). These processes of speaking and listening can be particularly difficult in second language (L2) academic settings because challenges arise from processing the language input or producing the language output in addition to the demands of processing the academic content itself (Gan, 2012). In academic settings, the most commonly observed classroom speaking and listening difficulties among non-native English speakers are comprehending lectures, presenting planned oral presentations, and participating in and leading classroom discussions (Evans & Morrison, 2011; Kim, 2006).

Further complicating these processes, learners' capabilities for L2 listening and speaking are often closely associated with each other (Bozorgian, 2012; Yoo & Manna, 2017). As Buck (2001) and Brown (2004) argue, language use represents an integration of skills rather than isolated capabilities and knowledge. Therefore, studies focusing on communication performance among EFL students often teach speaking and listening skills at the same time because a lack of either skill might lead to communication breakdowns (Rabab'ah, 2016). Metacognitive strategies pertaining to mutual understanding, such as paraphrasing the speaker, are also taught to develop the ability to maintain a two-way conversation (Fang, Cassim, Hsu, & Chen, 2018; Nakatani, 2005). Such studies have shown that training learners to be aware of their listening and speaking performance enhances their communication performance (Fang et al., 2018; Nakatani, 2005). However, to date, little research has examined listening and speaking together for adult Chinese professionals studying English, particularly in relation to metacognitive and self-assessment capabilities.

Speaking and listening in the Chinese EFL classroom

Classroom and teaching practices are unique to the cultures where they are practiced every day (Paris-Kidd & Barnett, 2011). For example, within many language classrooms in China, there are often separate segments of the lesson or distinct courses for each of the four skills where learners are asked to focus exclusively on listening, speaking, reading, or writing whereas other approaches, such as Communicative Language Teaching, unite these skills into a single class (Lee & VanPatten, 2003). Chinese culture may likewise impact students' approaches to their own learning and teachers' practices in facilitating learning (Paris-Kidd & Barnett, 2011). In fact, these classrooms are often designed to facilitate a transfer of knowledge from instructor to student, focusing on grammar for understanding. More specifically, in a survey of more than 800 Chinese university students in China, Maoying and Aiwu (2011) found that 42.5% of students felt that the key role of their instructor was to provide the knowledge in the classroom, acting as the "source of knowledge" (p. 83). In another study examining the Chinese EFL learning context, Zheng and Borge (2013) found that secondary school instructors who reported using more task-based and communicative pedagogy were likely to infuse their lessons with a strong focus on grammar despite stating that they preferred to focus on communicative activities, with their actions showing a preference for knowledge-transfer paradigms. Such classroom culture may shape how learners conceive of their language skills and their own abilities to self-assess their strengths and weaknesses in English.

With the majority of current studies with Chinese EFL learners focusing on secondary school and undergraduate university settings, few have examined the realities of EFL training courses for adults who are working as professionals. Yet, these learners are also re-engaging in their study of English language in order to expand their capabilities in English and attain their own personal and professional goals (Yang, 2006). As this population of professionals seeking further English proficiency grows, it is vital to expand our current examinations of listening and speaking in English with these professionals studying in mainland China.

A need for self-assessment

Within this study, self-assessment is broadly defined as a student's evaluation of their own capabilities. Such judgements may appear to be simple, but prior research indicates mixed results in learners' abilities to accurately represent their own abilities (Brantmeier, 2005, 2006; Brantmeier & Vanderplank, 2008; Ross, 1998; Suzuki, 2015). Yet, self-assessment is key to developing autonomous, life-long language learners. Specifically, as Little (2005) argues, learners must be involved in their own learning and assessment in order to develop their autonomous capacities, strengthening their abilities to manage their own language learning. Huang and Benson (2013) further support this understanding of self-assessment in their synthesis, constructing autonomy as a capacity that is concerned with learners engaging in understanding and controlling their own learning. Prior research has demonstrated positive results of training students to self-assess L2 listening and speaking. For example, students who received self-assessment training outperformed peers who did not experience the training on L2 listening and speaking (Mazloomi & Khabiri, 2016; Poehner, 2012; Shahrakipour, 2012; Sweet & Mack, 2017, March; Yoon & Lee, 2013). Understanding when and where students are able to self-assess may provide insights which could improve language learning outcomes through increased learner autonomy.

Examining self-assessment

Due to its potential benefits, self-assessment has been examined by many researchers in varied language contexts (Bachman & Palmer, 1989; Brantmeier, 2005, 2006; Brantmeier et al., 2012; Dolosic, 2018; Krausert 1991; LeBlanc & Painchaud, 1985; H. Liu & Brantmeier, 2019; Oskarsson, 1978; Ross, 1998). In many of these assessments, the Pearson correlation has been the preferred method for assessing the association among learners' performance and reported capabilities, with stronger correlations indicating that learners' own assessments and performance aligned (Brantmeier et al., 2012; Ross, 1998). Such relationships have demonstrated that many learners are able to accurately depict their own strengths and weaknesses with their own evaluations matching their performance (Brantmeier et al., 2012; Dolosic, 2018). However, to date, results appear mixed about whether learners are able to accurately depict their own strengths and weaknesses (Brantmeier, 2005, 2006). In seeking to understand these relationships more fully, investigations have demonstrated three areas which may impact on learners' abilities to accurately represent their own abilities in a foreign language. They are: (1) the instruments and questionnaires used, (2) individual and cultural traits, and (3) prior experiences using the language.

For example, when developing a self-assessment tool for a university Spanish language programme, Brantmeier (2005) sought to understand students' abilities to self-assess their reading in Spanish. Yet, her findings indicated that with descriptive items, students' self-ratings correlated with only some measures of L2 reading comprehension, and did not significantly discriminate among different student capabilities (Brantmeier, 2005). In a further investigation, using a computer-based test with the same population of students, student responses to descriptive self-assessment items did not align with their reading comprehension scores (Brantmeier, 2006). In seeking a more effective self-assessment tool, Brantmeier and Vanderplank (2008) built on the work of the DIALANG project (for a description of the project, see, Zhang & Thompson, 2004; for the project itself, see, https://dialangweb.lancaster.ac.uk/) to implement a criterion-referenced self-assessment that situated the learner in a specific context of language use. Findings for this self-assessment tool indicated that university students studying Spanish were able to accurately self-assess their reading comprehension when they were given this criterion-referenced self-assessment (Brantmeier & Vanderplank, 2008). Such results aligned with the findings of Ross' (1998) meta-analysis of self-assessment, indicating that more concrete and situated items lead to more accurate self-assessments. Continuing examinations of these criterion-referenced items, Brantmeier et al. (2012) expanded the instrument to include listening and writing, finding that students of university Spanish were able to accurately self-assess across these skills. Therefore, using criterion-referenced questionnaires may be the support that learners need to accurately depict their own strengths and weaknesses. Yet, there is a need to explore this finding across learning contexts in order to develop a full understanding of the role of the instrument in self-assessment accuracy.

In addition to the design of the instrument, learners' abilities to represent their own capabilities are often shaped by the learners' cultural and individual characteristics. For example, Hung, Samuelson, and Chen (2016) found that learners studying English in Taiwan were likely to rate themselves lower than their teacher would. The authors of this study suggest that these findings are tied to the cultural values in Taiwan which support humility as a positive attribute (Hung et al., 2016). Examining self-assessment alongside an interactive listening paradigm, Jingyan and Baldauf (2011), found that Chinese university students studying EFL demonstrated improvements on their interactive listening after a special classroom condition where learners completed an interactive

learning intervention, yet these students did not rate themselves more highly on self-assessments. Rather, their own ratings of their listening capabilities were nearly equal before and after training. The authors suggest that this may be explained by the strong preference for modesty that is present in Chinese society (Jingyan & Baldauf, 2011). Yet, with EFL reading comprehension and a criterion-referenced self-assessment questionnaire, Dolosic (2018) found that Chinese university students enrolled in an English for Academic Purposes class were able to accurately self-assess, particularly with narrative texts and multiple-choice tests. This apparent contradiction indicates a need for more careful examination of self-assessment across skills within this educational context. Individual differences have also been shown to shape L2 self-assessment. For example, MacIntyre, Noels, and Clément (1997) reported that learners with an L2 of French in Canada were less able to accurately represent their own strengths and weaknesses when they had higher levels of language anxiety. Proficiency in the language may also play a role as discovered by Heilenman (1990) in a study of more than 200 learners of French at a university in Canada which showed that learners at earlier stages in their language study often overestimate their abilities while those further into their programmes are likely to underestimate their abilities.

Alongside these personal characteristics, learners' experiences using their linguistic abilities have also demonstrated a role in accuracy of self-assessment. For example, with American high school students studying French in a specialized immersion summer camp, Dolosic, Brantmeier, Strube, and Hogrebe (2016) found that while students were unable to self-assess using a criterion-referenced self-assessment at the beginning of their experience, they were able to self-assess their French speaking accurately after the highly experiential and interactive language learning of the camp. Similarly, Suzuki (2015) investigated the experiential factors of Chinese university students studying Japanese, finding that factors such as time spent in Japan were closely tied to students' abilities to self-assess. These studies suggest that listening and speaking a language in everyday life may promote accurate self-assessment, even when learners are using criterion-referenced self-assessments.

While the studies reviewed above suggest potential for self-assessment as a language learning tool such studies are relatively sparse and the findings are not always consistent. Further research is needed to fully understand the use, accuracy and potential for self-assessment across varying contexts of language learning.

Research questions
1. What is the relationship between adult Chinese EFL learners' listening and speaking scores?
2. When measured with a criterion-referenced instrument, how do Chinese learners of English self-assess listening and speaking? Is there a relationship between learners' self-assessment and their subsequent performance on listening and speaking measures?

Materials and methods

Research design
This correlational, observational study used quantified measures of self-assessment, listening, and speaking in order to examine and understand the relationships among these variables within the context of adults experiencing a specialized English training

programme tailored to their needs. This study was conducted with a subset of data from a large-scale investigation examining a variety of linguistic and social psychology variables within the language learning context of this specialized English programme. Other aspects of the wider investigation are reported elsewhere (H. Liu & Brantmeier, 2019; Schultz, 2017). This paper focuses on understanding self-assessment, listening, and speaking within the context.

Participants and learning context

All students enrolled in one semester of the specialized EFL programme at a university in Northern China voluntarily participated in this study. This group contained eighty-two Chinese learners of English, aged 25-46, with balanced self-reported gender (forty-one men, forty-one women). All participants entered the programme with an intermediate level of English. The specialized English training programme is highly regarded within China and offers separate, intensive courses in reading, writing, listening, and speaking of English. In each class, approximately thirty learners have a single instructor who practices important tasks with students, such as listening to a recording for understanding. Using these tasks, teachers highlight difficulties and strategies to overcome challenges.

This programme prepares scholars of various disciplines for a year abroad at an English-speaking university, usually in the United Kingdom, Canada, Australia, or the United States. It prepares them to use English both in common daily tasks and discipline-specific discussions of research or advanced practice in fields such as medicine. In order to gain access to funding that supports their year abroad, the scholars are required to pass a test of English language across the skills of reading, writing, listening, and speaking. The programme assists participants in preparing for the test.

Instruments

Self-assessment data were collected using a criterion-referenced questionnaire consisting of items drawn from the *Reading, Writing, Listening and Speaking Self-assessment Grid* of the DIALANG project (as described earlier). It presented participants with statements about their language use such as "I can understand basic conversation if the speaker talks slowly" and asked them to respond using a five-point Likert scale ranging from *strongly disagree* to *strongly agree*. The self-assessment items were presented in the learners' L1 and divided into sections for listening and speaking. The Cronbach α for listening and speaking were high (listening Cronbach α = .95; speaking Cronbach α = .93), suggesting high reliability of this self-assessment questionnaire.

The language test used in this study was the same test that was used to determine whether students will have access to funding following their English language programme. It specifically tests the language knowledge and abilities needed for the exchange programme. Due to its place within the travel grant system in China, reliability and validity statistics of this test cannot be reported here. However, it was designed with care to provide a benchmark for scholars embarking on a year abroad.

The test includes speaking and listening sections. The listening section contained four sub-sections each of which focused on a different listening passage and contained   ten items. The listening passages are selected to examine a broad range of listening skills with one passage focusing on social needs and dialogues, another being a recorded news report, and two others that were lectures on education, living situations, or topics of general academic interest. Task types include true/false, short answer, fill-in-the-blank, and

summary completion. Scores on this assessment depict learners' accuracy in answering questions about the information presented through these recorded passages.

In the speaking section, examinees speak to two examiners: an observer and an interlocuter who both rate the student. This section includes three tasks. First, examinees speak about their work, family, hometown, and/or hobbies. Then, they are handed a cue card presenting a specific situation which they prepare (for one minute) and then respond to. Cards contain topics such as how to order food at a restaurant or discuss rent prices over the telephone. Finally, examinees respond to a complex social issue in China or abroad . Speaking for about three minutes they are expected to provide an opinion and support it with clear personal or academic examples. Topics cover areas such as the housing shortage in China or marriage and divorce rates in a community. The speaking performances are scored on a rubric that instructors are trained to use which include measures of both accuracy and fluency of English language use, focusing on the comprehensibility of the information that speakers share.

The listening and speaking task scores from both parts of the language test were compiled to provide an overarching understanding of participants' abilities to listen and speak in English.

Procedures
The data were collected in the final week of the semester. Participants signed a consent form. On a day separate from their courses, participants came to large lecture halls where they completed the self-assessment measures as well as the reading, writing, and listening portions of their language exam. For the speaking section, learners were distributed among raters who scored their performances on the three speaking tasks based on a rubric.

Analysis
First, descriptive statistics of key outcome variables were examined, and assumptions were checked. To examine the relationship between learners' listening and speaking performance scores, initial plotting of the listening and speaking performance scores were conducted, followed by a Pearson correlation analysis on actual listening and speaking performance. Another Pearson correlation was found among the self-assessment scores of listening and speaking. These scores were compared and analysed to understand differences between perceived abilities and performed abilities. All analyses were conducted with the open-source statistical program, R version 3.3.3 (R Core Team, 2017), using additional packages including psych (Reville, 2017), psychometric (Fletcher, 2010), and car (Fox & Weisberg, 2011). Figures were generated using GGPlot2 (Wickham, 2009).

Results

RQ1. What is the relationship between adult Chinese EFL learners' listening and speaking scores?
Learners within this particular context demonstrated a wide range of scores for both listening and speaking performance (see Table 1). Yet, listening performance demonstrated greater variability than speaking performance. Further, the two performance measures were significantly related, ($r = .41$, $p < .01$, see Figure 1), indicating a relationship between these skills for these participants.

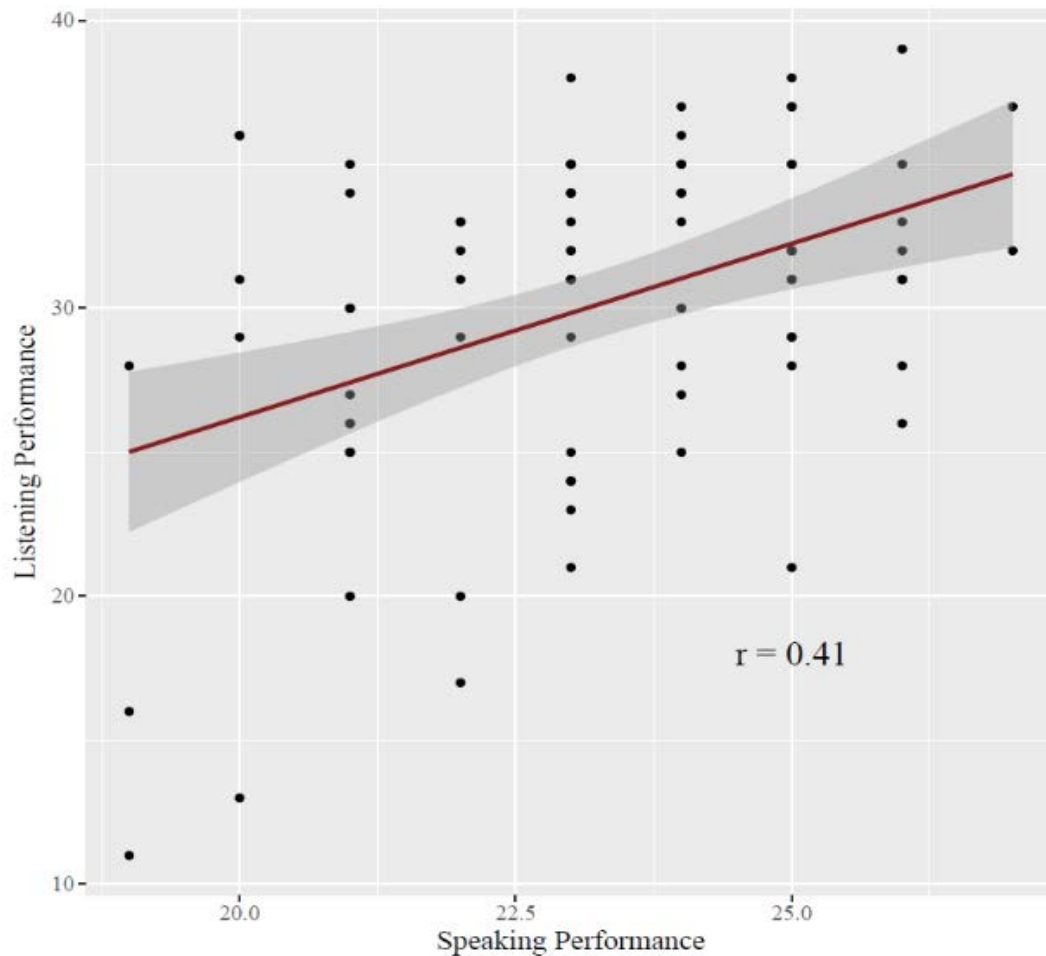Table 1. Descriptive and correlational statistics of listening and speaking performance

Figure 1. Association between listening and speaking performance

RQ2. When measured with a criterion referenced instrument measuring self-assessment, what are the speaking and listening ratings of Chinese learners of English? Is there a relationship between learners' self-assessment and their subsequent performance on listening and speaking?

Learners in this context averaged a neutral rating of their own performance; however, scores spanned the possible range (1-5), as can be seen in Table 2. Further, learners' self-assessment ratings correlated across skills ($r = .76$, $p < .01$). This demonstrates a perceived association of these skills which is larger than the association that performance measures demonstrated (see Figure 2). With a Fisher's Z transformation significance comparison, these correlations were examined, and correlations among listening self-assessment and speaking self-assessment were found to be greater than correlations among listening and speaking performance, with that difference being statistically significant ($p < .01$). However, no significant relationships were found between listening self-assessment and performance measures ($p > .05$) or speaking self-assessment and performance measures ($p > .05$), and in fact, the relationships were slightly negative (see

Figure 3). As such, students who performed worse on performance measures tended to assess themselves slightly more highly. However, this trend was not consistent enough to yield statistically significant results. Rather, this lack of statistically significant correlations demonstrates that students' self-evaluations did not consistently predict performance. Thus, any specific rating on a speaking or listening self-assessment did not associate with actual speaking or listening performance.

Table 2. Descriptive and correlational statistics of listening and speaking self-assessment

	Minimum	Maximum	Mean	Standard Deviation	Correlation
Listening SA	1.64	4.23	2.93	0.67	r = .76
Speaking SA	1.67	5.00	3.08	0.72	p < .01

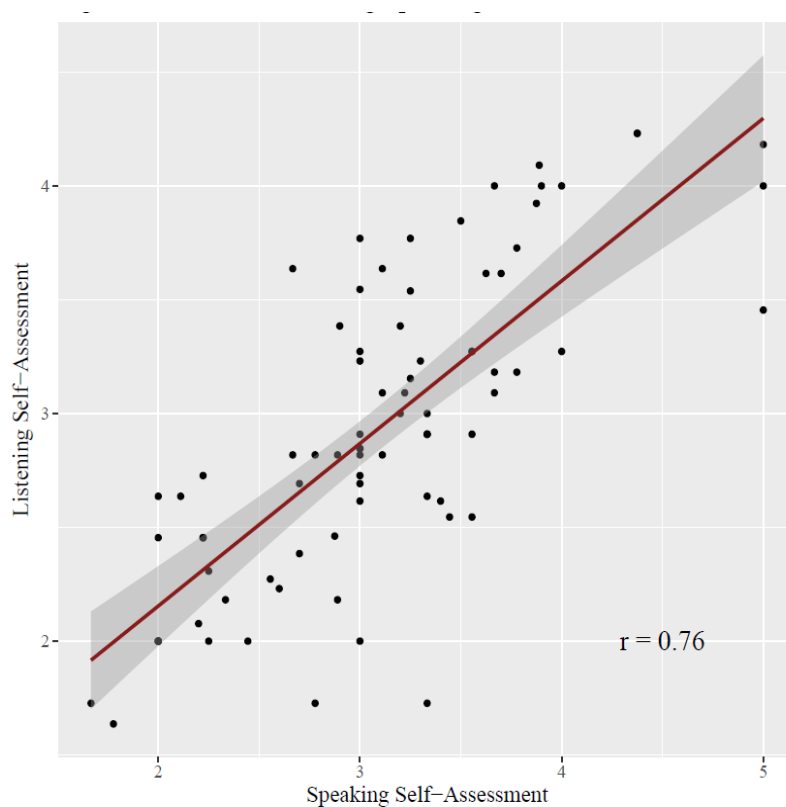Note: N = 82; SA is self-assessment

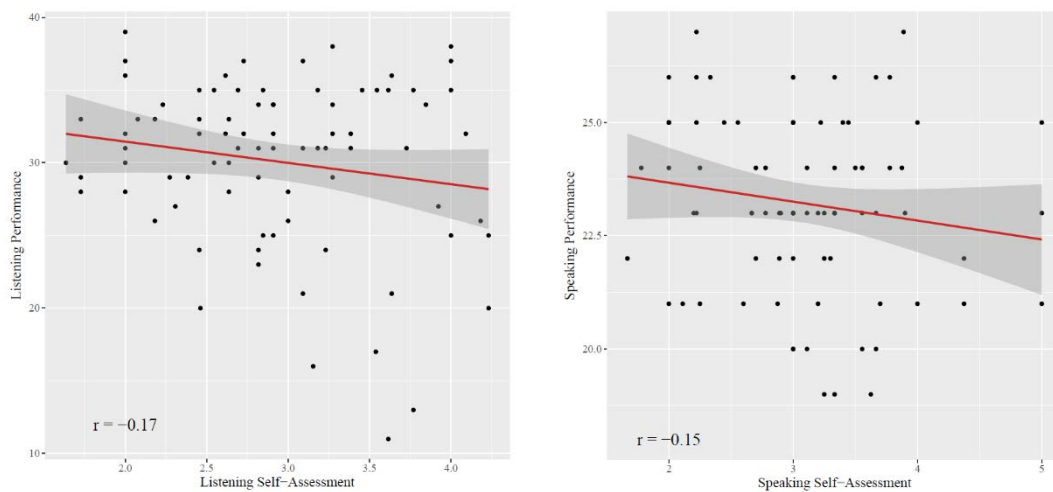Figure 2. Association among listening and speaking self-assessment

Figure 3. Relationships between speaking and listening self-assessment and performance

Discussion

Results demonstrate students were not able to accurately self-diagnose their strengths and weaknesses in speaking and listening as their self-ratings did not align with their performance. These findings both parallel and contradict previous findings due to the mixed results surrounding self-assessment in various language learning contexts (Brantmeier, 2005, 2006; Brantmeier et al., 2012; Dolosic et al., 2016; Krausert 1991; Oskarsson, 1978; Ross, 1998). Previous studies have, however, demonstrated that self-assessment is more accurate when learners are situated in a specific language use situation through activities or question items (Brantmeier & Vanderplank, 2008; Brantmeier et al., 2012; Ross, 1998) or when they have concrete language learning experiences (Dolosic et al., 2016; Suzuki, 2015). Yet, there are a number of factors (such as: the instruments and questionnaires used; individual and cultural traits; and prior experiences using the language) that shape accuracy in self-assessment (Dolosic, 2018). With the lack of a clear relationship among learners' self-assessments and performance in this study, it is possible that one of these factors has affected these self-assessments. This suggests the need to examine the multiple facets of self-assessment simultaneously and comprehensively.

Within this study, self-assessment was not found to align with learners' performance. One possible explanation for this result could be drawn from the learning context. More specifically, the commonly teacher-centred and grammar-focused ways in which Chinese EFL classrooms are typically structured (J. Y. Liu, Chang, Yang, & Sun, 2011; Maoying & Aiwu, 2011; Paris-Kidd & Barnett, 2011; Zheng & Borge, 2013) may impact these learners' abilities to self-assess. In the teacher-centred classroom where an instructor is seen primarily as the source of knowledge (Maoying & Aiwu, 2011), students may not be encouraged to reflect on their own abilities, instead allowing for their understanding of their own English language abilities to come from the instructor. In addition, teachers who focus on grammar may not be able to also provide time for students to work collaboratively in English, resulting in few interactive experiences, ultimately limiting experiences for learners to draw on when assessing their own abilities (Zheng & Borge, 2013). Yet, Gan (2012) argued that providing practice conditions that are interactive and more relevant to the actual conditions outside the classroom can help expose learners to

their own strengths and weaknesses. Alongside prior findings that learners with greater concrete experiences of language use are more able to accurately self-assess (Dolosic et al., 2016; Suzuki, 2015), these results suggest that these Chinese EFL learners may be less accurate in self-assessment because they have not had such reflective or interactive experiences in their learning of listening and speaking.

Another explanation for these results could emerge from the culture that surrounds these students' learning experiences. As researchers have indicated previously in studies of self-assessment, within China there is a preference for modesty (Hung et al., 2016; Jingyan & Baldauf, 2011). In prior studies of Chinese learners, individuals have appeared to underrate their own performances and abilities consistently. In this way, these cultural values may shape how some learners approach self-assessment of their strengths and weaknesses. However, the results of this study do not indicate a significant trend where all learners underrate their own performance. Rather, the wide dispersion of data suggests that while some learners acted humbly in rating their own performance, this was not the overarching trend. These results align with the complexity of self-assessment ratings found in varied language learning contexts (Brantmeier, 2005, 2006; Brantmeier et al., 2012; Dolosic, 2018).

The findings in this study further affirm the relationship between listening and speaking performance found in prior studies (Fang et al., 2018; Nakatani, 2005; Rabab'ah, 2016). Yet, the results of this study also indicated that the association among learners' self-ratings did not match up with the association among actual performance abilities. Specifically, students' self-ratings of listening and speaking were more highly correlated than their actual performance measures of listening and speaking, possibly demonstrating that learners believe that their skills are more balanced than their performance indicates. Such a discrepancy further highlights the need for these learners to develop a better understanding of their own strengths and weaknesses (Brantmeier, 2006; Brantmeier et al., 2012).

With reference to Chinese EFL classrooms the findings of this study suggest that learners need greater interactive opportunities to explore their own individual strengths and weaknesses when encountering language so that they will be able to develop capabilities to become autonomous, life-long language learners.

Limitations

The use of composite exam scores to represent performance prevented the investigation of test method effects. This practice is consistent with prior research in self-assessment. However, it would be interesting in future research to investigate this aspect of the complex relationships between self-assessment and performance. Further, the sample was relatively small so more and larger investigations are needed to expand understanding of this growing population of adult EFL learners.

Conclusion

Findings of this study clearly demonstrate that these Chinese EFL adult learners were unable to accurately depict the strengths and weaknesses of their English listening and speaking abilities. These results, taken within the reality of Chinese classroom culture, indicate a need for greater understanding of the associations among Chinese EFL self-assessment abilities and listening and speaking. As learning experience has been shown to shape learners' perceptions of their weaknesses and strengths in speaking and listening, the pedagogical implications for practitioners are: (1) to create learning conditions in

which learners can practice and become aware of their language gap, thus potentially better self-assessing their abilities, and (2) to integrate formative assessment tools such as self-assessment questionnaires to help learners reflect on their own speaking and listening abilities throughout the learning process. Future studies should examine self-assessment across different instructional contexts and investigate the effectiveness of specialized self-assessment training. Further research in these areas will contribute to the development of self-assessment as a tool for developing learner autonomy and supporting life-long learning.

About the authors

Haley Dolosic recently completed her PhD in applied linguistics in education in the Department of Education at Washington University in St. Louis. Her current research interests include self-assessment across diverse linguistic backgrounds and advanced research methodology in applied linguistics.

Wei-Chieh Fang has an interdisciplinary background in applied linguistics and cognitive psychology. His recent research explores the potential of using smartphones to facilitate communicative tasks and the development of strategic competence in speaking.

Cindy Brantmeier is a professor of applied linguistics and education and Director of Applied Linguistics at Washington University in St. Louis. She is also the co-editor of *Reading in a Foreign Language* and was named Distinguished Visiting Professor, Northeast Normal University, Changchun, China. Her published work consists of experiments that examine variables involved in second language reading, language research methodology, and language testing and assessment.

Michael Strube is a professor of psychological and brain sciences at Washington University in St. Louis. His research interests include self and identity, statistics, and research design.

Yanming Gao is an associate professor at the Training Centre of the Ministry of Education for Studying Overseas at Northeast Normal University in Changchun, China. Her current research interests include pragmatic failure of adult Chinese learners of English, and EFL teaching and learning in applied linguistics.

References

Bachman, L. F., & Palmer, A. S. (1989). The construct validation of self-ratings of communicative language ability. *Language Testing, 6*(1), 14-29. https://doi.org/10.1177/026553228900600104

Bozorgian, H. (2012). The relationship between lListening and other language skills in international English language testing system. *Theory and Practice in Language Studies, 2*(4), 657-663. https://doi.org/10.4304/tpls.2.4.657-663

Brantmeier, C. (2005). Nonlinguistic variables in advanced second language reading: Learners' self-assessment and enjoyment. *Foreign Language Annals, 38*(4), 494-504. https://doi.org/10.1111/j.1944-9720.2005.tb02516.x

Brantmeier, C. (2006). Advanced L2 learners and reading placement: Self-assessment, CBT, and subsequent performance. *System: An International Journal of Educational Technology and Applied Linguistics, 34*(1), 15-35. https://doi.org/10.1016/j.system.2005.08.004

Brantmeier, C., & Vanderplank, R. (2008). Descriptive and criterion-referenced self-assessment with L2 readers. *System, 36*(3), 456-477. https://doi.org/10.1016/j.system.2008.03.001

Brantmeier, C., Vanderplank, R., & Strube, M. (2012). What about me?: Individual self-assessment by skill and level of language instruction. *System, 40*(1), 144-160. https://doi.org/10.1016/j.system.2012.01.003

Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Boston: Allyn & Bacon.

Buck, G. (2001). *Assessing listening*. Cambridge, UK: Cambridge University Press.

Cheng, L. (2008). The key to success: English language testing in China. *Language Testing, 25*(1), 15-37. https://doi.org/10.1177/0265532207083743

Dolosic, H. (2018). An examination of self-assessment and interconnected facets of second language reading. *Reading in a Foreign Language, 30*(2), 189-208.

Dolosic, H., Brantmeier, C., Strube, M., & Hogrebe, M. C. (2016). Living language: Self-assessment, oral production, and domestic immersion. *Foreign Language Annals, 49*(2), 302-316. https://doi.org/10.1111/flan.12191

Evans, S., & Morrison, B. (2011). Meeting the challenges of English-medium higher education: The first-year experience in Hong Kong. *English for Specific Purposes, 30*(3), 198-208. https://doi.org/10.1016/j.esp.2011.01.001

Fang, W.-C., Cassim, F. A. K., Hsu, C.-N., & Chen, N.-S. (2018). Effects of reciprocal peer feedback on EFL learners' communication strategy use and oral communication performance. *Smart Learning Environments, 5*(1), 11. https://doi.org/10.1186/s40561-018-0061-2

Fletcher, T. (2010). psychometric: Applied psychometric theory [Computer Software]. Retrieved from http://www.cran.r-project.org/

Fox, J., & Weisberg, S. (2011). *An R Companion to Applied Regression,* (2nd ed.). Thousand Oaks, CA: Sage.

Gan, Z. (2012). Understanding L2 speaking problems: Implications for ESL curriculum development in a teacher training institution in Hong Kong. *Australian Journal of Teacher Education, 37*(1), 43-59. https://doi.org/10.14221/ajte.2012v37n1.4

Heilenman, K. L. (1990). Self-assessment of second language ability: The role of response effects. *Language Testing, 7*(2), 174-201. https://doi.org/10.1177/026553229000700204

Huang, J., & Benson, P. (2013). Autonomy, agency and identity in foreign and second language education. *Chinese Journal of Applied Linguistics, 36*(1), 7-28. https://doi.org/10.1515/cjal-2013-0002

Hughes, R., & Reed, B. S. (2016). *Teaching and researching speaking*. New York, NY: Taylor & Francis.

Hung, Y. J., Samuelson, B. L., & Chen, S. C. (2016). Relationships between peer-and self-assessment and teacher assessment of young EFL learners' oral presentations. In M. Nicolov (Ed.), *Assessing young learners of English: Global and local perspectives* (pp. 317-338). Berlin: Springer International Publishing.

Jingyan, C., & Baldauf, R., B. (2011). Can EFL interactive listening be validly assessed? *Chinese Journal of Applied Linguistics, 34*(4), 31-46. https://doi.org/10.1515/cjal.2011.032

Kim, S. (2006). Academic oral communication needs of East Asian international graduate student in non-science and non-engineering fields. *English for Specific Purposes, 25*, 479-489. https://doi.org/10.1016/j.esp.2005.10.001

Krausert , S. R. (1991). *Determining the usefulness of self-assessment of foreign language skills: post-secondary ESL students' placement contribution* (Unpublished doctoral dissertation), University of Southern Carolina, Columbia, SC.

Lai, E. (2001). Teaching English as a private enterprise in China. *English Today, 17*(2), 32-36. https://doi.org/10.1017/S0266078401002048

LeBlanc, R., & Painchaud, G. (1985). Self-assesment as a second language placement instrument. *TESOL Quarterly, 19*(4), 673-687.

Lee, J., & VanPatten, B. (2003). *Making communicative language teaching happen*. Blacklick, OH: McGraw-Hill Education.

Li, X., & Cutting, J. (2011). Rote learning in Chinese culture: Reflecting active confucian-based memory strategies. In L. Jin & M. Cortazzi (Eds.), *Researching Chinese learners: Skills, perceptions and intercultural adaptations* (pp. 21-42). Basingstoke, UK: Palgrave MacMillan.

Little, D. (2005). The Common European Framework and the European Language Portfolio: Involving learners and their judgements in the assessment process. *22*(3), 321-336. https://doi.org/10.1191/0265532205lt311oa

Little, D. (2009). Language learner autonomy and the European Language Portfolio: Two L2 English examples. *Language Teaching, 42*(2), 222-233. https://doi.org/10.1017/S0261444808005636

Liu, H., & Brantmeier, C. (2019). "I know English": Self-assessment of foreign language reading and writing abilities among young Chinese learners of English. *System, 80*, 60-72. https://doi.org/10.1016/j.system.2018.10.013

Liu, J. Y., Chang, Y. J., Yang, F. Y., & Sun, Y. C. (2011). Is what I need what I want? Reconceptualising college students' needs in English courses for general and specific/academic purposes. *Journal of English for Academic Purposes, 10*(4), 271-280. https://doi.org/10.1016/j.jeap.2011.09.002

Liu, M., & Jackson, J. (2008). An exploration of Chinese EFL learners' unwillingness to communicate and foreign language anxiety. *The Modern Language Journal, 92*(1), 71-86. https://doi.org/10.1111/j.1540-4781.2008.00687

MacIntyre, P. D., Noels, K. A., & Clément, R. (1997). Biases in self-ratings of second language proficiency: The role of language anxiety. *Language Learning, 47*(2), 265-287.

Maoying, X., & Aiwu, J. (2011). Comparison between the students' perceived and expected behavior of college English teachers. *Chinese Journal of Applied Linguistics, 34*(4), 72-88. https://doi.org/10.1515/CJAL.2011.035

Mazloomi, S., & Khabiri, M. (2016). Diagnostic assessment of writing through dynamic self-assessment. *International Journal of English Linguistics, 6*(6), 19-31. https://doi.org/10.5539/ijel.v6n6p19

Nakatani, Y. (2005). The effects of awareness-raising training on oral communication strategy use. *The Modern Language Journal, 89*(1), 76-91.

Nguyen, L. T. C., & Gu, Y. (2013). Strategy-based instruction: A learner-focused approach to developing learner autonomy. *Language Teaching Research, 17*(1), 9-30. https://doi.org/10.1177/1362168812457528

Oskarsson, M. (1978). *Approaches to self-assessment in foreign language learning*. Strasbourg, France: Council for Cultural Cooperation.

Paris-Kidd, H., & Barnett, J. (2011). Cultures of learning and student participation: Chinese learners in a multicultural English class in Australia. . In L. Jin & M. Cortazzi (Eds.), *Researching Chinese learners: Skills, perceptions and intercultural adaptions* (pp. 169-187). Basingstoke, UK: Palgrave Macmillan.

Poehner, M. E. (2012). The zone of proximal development and the genesis of self-assessment. *The Modern Language Journal, 96*(4), 610-622.

R Core Team. (2017). R: A language and environment for statistical computing [Computer Software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.cran.R-project.org/

Rabab'ah, G. (2016). The effect of communication strategy training on the development of EFL learners' strategic competence and oral communicative ability. *Journal of Psycholinguistic Research, 45*(3), 625-651.

Reville, W. (2017). psych: Procedures for personality and psychological research [Computer Software]. Retrieved from https://www.cran.R-project.org/

Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing, 15*(1), 1-20. https://doi.org/10.1177/026553229801500101

Rost, M. (2002). *Teaching and researching listening*. Harlow: Longman.

Schultz, L. M. (2017). Affect with Chinese learners of English: Enjoyment, self-perception, self-assessment, and abilities across levels of language learning 1. *Quarterly Journal of Chinese Studies, 5*(2), 65-81.

Shahrakipour, H. (2012). On the impact of self-assessment on EFL learners' receptive skills performance. *International Research Journal of Arts and Humanities, 40*(40), 1-13.

Suzuki, Y. (2015). Self-assessment of Japanese as a second language: The role of experiences in the naturalistic acquisition. *Language Testing, 32*(1), 65-81. https://doi.org/10.1177/0265532214541885

Sweet, G., & Mack, S. (2017, March). *Self-assessment and learner agency: A new approach.* Paper presented at the the meeting of the American Association for Applied Linguistics Portland, OR.

Wickham, H. (2009). Ggplot2: Elegant graphics for data analysis [Computer Software]. New York, NY: Springer-Verlag. Retrieved from http://www.cran.r-project.org

Wu, S. (2016). *The use of L1 cognitive resources in L2 reading by Chinese EFL learners*. London: Routledge.

Yang, J. (2006). Learners and users of English in China. *English Today, 22*(2), 3-10.

Yoo, H., & Manna, V. F. (2017). Measuring English language workplace proficiency across subgroups: Using CFA models to validate test score interpretation. *Language Testing, 34*(1), 101-126.

Yoon, E., & Lee, H. K. (2013). Do effects of self-assessment differ by L2 language level? A case of Korean learners of English. *The Asia-Pacific Education Researcher, 22*(4), 731-739.

Zhang, S., & Thompson, N. (2004). DIALANG: a diagnostic language assessment system. *The Canadian Modern Language Review/La Revue Canadienne des Langues Vivantes, 61*(2), 290-293. https://doi.org/10.1353/cml.2005.0011

Zheng, X., & Borge, S. (2013). Task-based learning and teaching in China: Secondary school teachers' beliefs and practices. *Language Teaching Research, 18*(2), 205-221. https://doi.org/10.1177/1362168813505941