Validation of the Self-assessment Scale for Reading Competence in the Online Diagnostic Assessment

Lanhui Zhu

Laolong Normal School, Heyuan Polytechnic, Heyuan, China

124478194@qq.com

Abstract. The self-assessment scale is an important instrument for English learners to self-diagnose and self-assess their language competence and is of great value in the diagnostic Assessment. Validation is a prerequisite and guarantee for its effective application. This research concentrates on the self-assessment of reading competence and reading strategy and adapt scales from the descriptors in the China's Standards of English Language Ability(CSE). The empirical findings show that the two scales have good construct validity. The feedback from the self-assessment provide diagnostic information about reading ability and reading strategy for learners. This research also proposes suggestions for improvement and application of the reading competence scale in the diagnostic Assessment and teaching.

Keywords: Self-assessment Scale; Reading Competence; Reading Strategy; Diagnostic Assessment.

1 Introduction

Self-assessment is a method for learners to reflect on their own learning, it serves to assist learners to discover their strengths and weaknesses, and make them aware of what needs to be improved[1]. It aims to motivate learners' awareness and autonomy in learning [2-3]. Self-assessment is the primary means for learners to enhance their awareness of language proficiency[4]. The validity of the scales is crucial in determining whether the self-assessment can promote learning. Teachers often concentrate on evaluating students' reading competence while paying less attention to their awareness of language proficiency in English reading teaching.

Based on the reading competence scale of the China's Standards of English Language Ability (CSE) [5], this study adapts self-assessment scales for reading competence and reading strategy and verifies their construct validity in an online reading diagnostic assessment. Furthermore, suggestions are also proposed for the future use and feedback of self-assessment scales in diagnostic assessment and teaching.

M. Yu et al. (eds.), *Proceedings of the 2024 5th International Conference on Big Data and Informatization Education (ICBDIE 2024)*, Advances in Intelligent Systems Research 182,
https://doi.org/10.2991/978-94-6463-417-4_15

2 Literature Review

2.1 Self-assessment and Diagnostic Assessment

Since the 1970s, there has been a growing research interest in self-assessment in the field of language learning. Current research concentrates on the reliability and validity of learners' self-assessment and factors influencing its accuracy. There is limited research on the validity of assessment scales. Compared to other external assessment methods such as exams, self-assessment has lots of advantages. It reduces the burden on teachers, enhances students' awareness of autonomous learning, enriches learning evaluation methods, and promotes learning [6]. Studies have shown that implementing of self-assessment in language testing can enhance learners' self-awareness of language proficiency, provide diagnostic information for language learning, and improve learners' self-awareness[7-8] , assisting them to understand their learning goals, enhance their autonomy in learning[9-10].

Research has gradually focused on the relationship between self-assessment and language proficiency testing, thereby verifying the validity of the self-assessment tools[11]. Dolosic found a significant correlation between learners' self-assessment and their actual reading performance[12]. Increasing the amount of reading can improve the accuracy of self-assessment, and the more specific the descriptors in the self-assessment scale, the higher the accuracy of learners' self-assessment [13-14].

As it's an alternative form of assessment, self-assessment is not widely accepted in large-scale and high-stakes language tests. However, there is still considerable space for exploration in low-stakes diagnostic assessments. It can diagnose learners' language proficiency, identify their strengths, and reflect on their weakness based on students' previous study[15-16]. It helps teachers to reconsider teaching plans, revise or adjust teaching procedures. It aims to "determine the causes of learning problems and propose remedial action"[17]. The diagnostic assessment scale acts as a bridge between assessment and learning, with great potential for assisting learning [18]. Self-assessment in diagnostic assessments can truly promote learning through assessment.

Currently, researches on diagnostic assessments mainly concentrate on cognitive diagnostic models[19-22], while studies of self-assessment in diagnostic assessments are relatively scare. Therefore, it is necessary to gain more valuable feedback information for diagnostic assessments through self-assessment, helping learners plan their learning objectives and improve their language proficiency.

2.2 Self-assessment of Reading Competence

There should be clear content and dimensions for self-assessment in diagnostic assessments, which is the primary source of the construct validity of the scales. Scholars have different opinions on the definition of reading competence. The multidimensional view divides reading competence into various elements such as recognition ability, comprehension ability, and inference ability, suggesting that reading requires the simultaneous application of various abilities. Currently, this viewpoint is widely accepted and has received theoretical, empirical, and instructional support. Language

testing experts and researchers also demonstrate the distinctiveness of reading competence through statistical methods such as factor analysis and correlation analysis.

Davis measured five abilities when students read texts: understanding the meanings of words, identifying author's purpose/attitude/tone, making inferences, answering questions, and clarifying the organization of the text. Guthrie and Kirsch extracted two dimensions of English reading competence by factor analysis: understanding the meaning of the text and locating information. Other studies have explored second language reading competence through correlation analysis and found that vocabulary knowledge, grammar knowledge, and reading strategies are important components of reading competence. It's suitable to compile teaching curriculum and activities to divide reading competence into different sub-abilities. Teachers can train students' various abilities to improve their overall reading competence according to individual differences.

CSE provides a broad definition and description of language competences, which is useful for self-assessment of reading competence. The CSE scale defines reading competence as the ability for language learners or users to construct meaning when reading written materials, using various knowledge (linguistic and non-linguistic) and strategies. It is divided into into three dimensions: the ability to identify and extract written information, the ability to summarize and analyze written information, and the ability to critically evaluate written information. There is a hierarchical progression in the three dimensions, which are consistent with Bloom and Anderson's cognitive taxonomy. The ability to identify and extract written information refers to the capacity to accurately recognize and reproduce specific information based on the reading materials. The ability to summarize and analyze written information refers to the capacity to grasp the overall elements of the reading material and make reasonable inferences and predictions based on comparisons and summarization of the materials. The ability to critically evaluate written information refers to the capacity to reflect on and evaluate the content, form, style, and intention of the reading materials by integrating prior knowledge with the reading materials.

Reading strategies are also one of the components of the CSE reading competence scale. Strategic competence consists of planning, implementing, evaluating and remedying approaches.

Self-assessment of reading competence is conducted before reading tests, and self-assessment of reading strategies is conducted after reading tests.

3 Research Design

3.1 Research Questions

The following two research questions are put forward:

1)What is the validity of the self-assessment scale for reading competence?

2)What is the validity of the self-assessment scale for reading strategies?

3.2 Design of the Self-Assessment Scales

This study is based on the online diagnostic assessment platform "Youzhenxue (University Edition)" designed by the Foreign Language Teaching and Research Press. The platform consists of four levels of tests, corresponding to the level four, level five, level six, and level seven of the CSE. The descriptors of the self-assessment scale for reading competence are derived from the reading competence scale in the CSE Level Four, totaling 25 items. The descriptors of the reading strategy self-assessment scale are derived from the reading strategy scale in the CSE Level Four, totaling 11 items.

The self-assessment scales use the phrase "I am able to" for descriptors and adopt the Likert 5-point scale (1-Strongly Disagree; 2-Disagree; 3-Neither Agree nor Disagree; 4-Agree; 5-Strongly Agree).

In terms of research procedure, participants first finish the self-assessment for reading competence, then log into the "Youzhenxue" platform (University Edition) to complete the reading diagnostic tests, and finally participate in the self-assessment for reading strategies.

3.3 Participants and Data Collection

Convenient random sampling was used in this study, with a sample of 280 first-year students from a vocational college in Guangdong Province, China. A total of 236 valid questionnaires were collected. Participants completed the self-assessment scales for reading c and reading strategies before and after the diagnostic reading tests.

Teachers distributed the questionnaires online and instructed the students in the classroom. The instructions included the reading competence scale, research introduction, and an explanation of the meaning of the response options "1-5." This research adopted statistical software SPSS 26.0 and AMOS 26.0 for data analysis, including confirmatory factor analysis and structural equation modeling.

4 Results

4.1 Self-assessment of Reading Competence

Table 1. the descriptive statistics for the self-assessment of reading competence

N	Mean	Std. Error of Mean	Media	Std. Deviation	Skew-ness	Std. Error of Skew-ness	Kurto-sis	Std. Error of Kurto-sis	Mini-mum	Maxi-mum	Percentiles		
											25	50	75
236	46.33	0.8208	48	12.61	-0.173	0.158	-0.046	0.316	16	79	38	48	54.75

Table 1 is the descriptive statistics for the self-assessment of reading competence. The scores rang from a minimum of 16 to a maximum of 79, with a mean of 46.33 and a

standard deviation of 12.61. Table 2 is the descriptive statistics for the self-assessment of reading competence subscale. The Cronbach's alpha reliability coefficient is 0.966, indicating high internal consistency. Further examination of the construct validity is conducted through confirmatory factor analysis. The results show good fit indices (CMIN/DF 1.998, GFI 0.911, AGFI 0.877, CFI 0.971, TLI 0.964, RMSEA 0.065). These findings indicate that the self-assessment scale for reading ability has satisfactory validity and can reliably reflect the dimensions of the reading competence, including the abilities of information identification and extraction, summarization and analysis, and critical evaluation.

Table 2. the descriptive statistics for the self-assessment of reading competence subscale.

	Mean	Std. Deviation	Skewness		Kurtosis	
			Statistic	Std. Error	Statistic	Std. Error
S14	3.2839	0.92248	-0.201	0.158	-0.034	0.316
S07	3.1695	0.95699	-0.052	0.158	-0.291	0.316
S01	3.1229	0.98378	0.048	0.158	-0.349	0.316
S12	2.9958	0.92941	-0.088	0.158	-0.098	0.316
S15	2.9958	0.99144	0.114	0.158	-0.367	0.316
S10	2.8983	0.95774	0.088	0.158	-0.071	0.316
S09	2.8898	0.96124	-0.038	0.158	-0.356	0.316
S11	2.8475	0.90509	-0.110	0.158	-0.285	0.316
S16	2.8347	0.94656	0.093	0.158	-0.284	0.316
S08	2.8136	1.01015	0.057	0.158	-0.306	0.316
S06	2.7966	0.94585	-0.131	0.158	-0.407	0.316
S13	2.7924	0.91052	0.014	0.158	-0.217	0.316
S05	2.7627	0.97787	0.023	0.158	-0.355	0.316
S04	2.7458	1.03296	0.036	0.158	-0.381	0.316
S02	2.7246	1.02126	0.211	0.158	-0.230	0.316
S03	2.6568	1.02147	0.197	0.158	-0.361	0.316

4.2 Self-assessment of Reading Strategies

Table 3 is the descriptive statistics for the self-assessment of reading strategies, arranged in descending order of the mean scores for each item. The self-assessment of reading strategies was conducted after participants completed the diagnostic reading test. The internal reliability the scale is 0.958.

Table 3. the descriptive statistics for the self-assessment of reading strategies

	Mean	Std. Deviation	Skewness		Kurtosis	
			Statistic	Std. Error	Statistic	Std. Error
Q07	3.1144	1.02295	-0.111	0.158	-0.302	0.316
Q05	3.0847	1.03205	-0.101	0.158	-0.276	0.316
Q01	3.0254	0.97598	0.170	0.158	-0.136	0.316
Q11	2.9915	0.91787	-0.183	0.158	-0.057	0.316
Q02	2.9746	0.95393	0.110	0.158	-0.132	0.316
Q03	2.9534	0.96423	-0.079	0.158	-0.144	0.316
Q06	2.9449	0.97257	0.167	0.158	-0.058	0.316
Q08	2.9280	0.98882	0.039	0.158	-0.142	0.316
Q10	2.8856	0.99341	-0.031	0.158	-0.253	0.316
Q09	2.8686	0.97398	0.071	0.158	-0.090	0.316
Q04	2.8136	0.97150	0.101	0.158	-0.061	0.316

The most frequently used strategies by students are Q07 (I can remember the reading purpose when I read a text) and Q05 (I can identify important information in a text by skimming, scanning, and skipping skills.). The least frequently used reading strategies are Q09 (I can analyse the writing style such as comparison to understand a text better) and Q04 (I can understand the meanings of words by word-formation).

The 11 descriptors in Table 3 are divided into three types of strategies: planning, implementing, evaluation and remediation. Further analysis reveals scores of these three strategies are: planning(mean=2.9816, SD=0.86494), evaluation and remediation (mean=2.9576, SD=0.86200), and implementing (mean=2.9506, SD=0.89332). This suggests that students tend to use planning strategies more frequently than other strategies during the reading process.

To further validate the dimensions of the self-assessment scale for reading strategies, a structural equation modeling approach is adopted. Given the inter-correlations among the three types of strategies, a path analysis is used to establish a three-factor structural model for reading strategies, as shown in Figure 1. The planning strategies(gs) consists of 3 descriptors(Q1-Q3); The implementing strategies(zx) consists of 3 descriptors(Q4-Q6); the evaluation and remediation strategies(pg) consists of 5descriptors(Q7-Q11). The model shows good fit indices (CMIN/DF 1.817, GFI 0.948>0.9, AGFI 0.916>0.9, CFI 0.985, TLI 0.980, RMSEA 0.059<0.08).

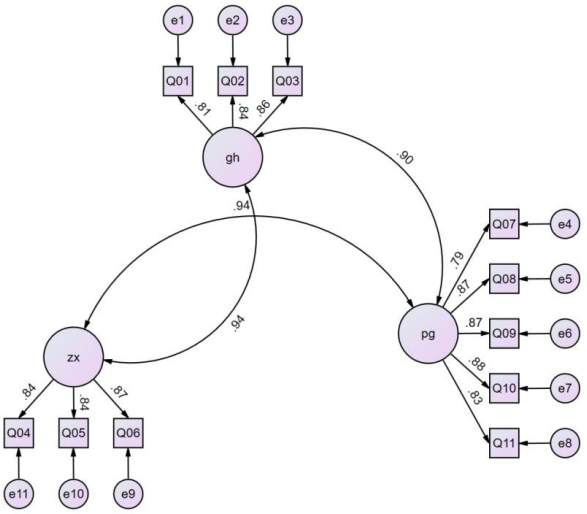

Fig. 1. Factor structural model for reading strategies.

The examination process verifies the convergent validity (AVE) and composite reliability (CR) of each dimension, as shown in Table 4. According to the criteria, the AVE are all greater than 0.5, and the CR are all greater than 0.7. This indicates that the self-assessment scale for reading strategies shows good convergent validity and composite reliability.

Table 4. the convergent validity (AVE) and composite reliability (CR) of each dimension

			Estimate	AVE	CR
Q01	<-	gh	0.808		
Q02	<-	gh	0.837	0.6983	0.8794
Q03	<-	gh	0.861		
Q04	<-	zx	0.840		
Q05	<-	zx	0.841	0.7262	0.8883
Q06	<-	zx	0.875		
Q07	<-	pg	0.788		
Q08	<-	pg	0.873		
Q09	<-	pg	0.869	0.7049	0.9227
Q10	<-	pg	0.877		
Q11	<-	pg	0.831		

5 Conclusions

This study focused on self-assessment in reading diagnostic assessment and adapted the self-assessment scales for reading competence and reading strategies from the

CSE Reading competence Scale. The results show that self-assessment is an effective means of evaluating reading competence and a valuable method for students' self-assessment in diagnostic assessment. The construct validity of reading competence and reading strategy scales are satisfactory, reflecting the dimensions of reading ability, including information identification and extraction, summarization and analysis, and critical evaluation, and the dimensions of reading strategy.

Moreover, self-assessment can be effectively achieved by combining the advantages of self-assessment scales with online diagnostic assessment, providing strong evidence for the accurate feedback in diagnostic assessments. This study offers insights into reading teaching. In reading teaching, teachers should provide learners with materials suitable for their current proficiency levels, design various reading tasks, and encourage students to take responsibility for their own reading behaviors, cultivating their autonomy in learning. It's a long-term process to improve reading competence by accumulation, online reading diagnostic tests, and self-assessment, promoting development through reading and learning through evaluation. From the perspective of learners, they can benefit from self-assessment by motivating their awareness of learning and strengthening the effective of strategy utilization.

In the future research, we will further explore the effect of self-assessment in learners' awareness of reading competence and setting learning objectives, providing insights and considerations for related studies.

Acknowledgment

This research was substantially supported by the Research Project of Education and Science in Guangdong Province "13th Five-Year Plan" in 2020: A Study on the Pre-service and In-service Training System for Commissioned Normal School Students based on Outcome-Based Education (OBE) -A Case Study of Heyuan Polytechnic (2020GXJK548), and the Research Project of the 4th Council of Guangdong Vocational and Technical Education Association in 2023-2024 (202212G190).

References

1. Hanrahan, S. J., Isaacs, G. Assessment and peer-assessment: The students' views[J]. Higher Education Research and Development, 2001, 20(1):53-70.
2. Oscarson, M. Self-assessment of language proficiency: Rationale and applications[J]. Language Testing, 1989, (1): 1-13.
3. Butler, Y. G., Lee, J. The effects of assessment among young learners of English[J]. Language Testing, 2010, 27(1): 854-867.
4. Pan Mingwei. Developing and validating the self-assessment scales in an online diagnostic test of English writing[J]. Foreign Language Education in China,2019,2(04):33-41.
5. Liu Jian-da. China's Standards of English Language Ability[M], Bei jing: Higher Education Press,2018.
6. Yu-ju Hung. Bridging assessment and achievement: Repeated practice of self-assessment in college English classes in Taiwan[J]. Assessment & Evaluation in Higher Education, 2019, 44(8): 1191-1208.

7. Alderson, J. C. Diagnosing Foreign Language Proficiency: The Interface Between Learning and Assessment [M]. New York: Continuum,2005.
8. Liu Jian-da. Self-assessment of English Writing Skills by Chinese University Students[J], Modern Foreign Languages,2002,25(03):241-249.
9. Suzuki, Y. Self-assessment of Japanese as a second language: The role of experiences in the naturalistic acquisition[J]. Language Testing, 2015, 32(1): 63-81.
10. Jamrus, M. H. M. & A. B. Razali. Using self-assessment as a tool for English language learning[J]. English Language Teach-ing, 2019(11): 64-73.
11. Ross, J. A. The reliability, validity and utility of self-assessment[J]. Practical Assessment, Research, and Evaluation, 2006(10): 1-13.
12. Dolosic, H. An examination of self-assessment and interconnected facets of second language reading[J]. Reading in a Foreign Lan-guage, 2018(2): 189-208.
13. Ross, S. Self-assessment in second language testing: A meta-a-nalysis and analysis of experiential factors[J]. Language Testing, 1998(1): 1-20.
14. Li, M, X. Zhang. A meta-analysis of self-assessment and lan-guage performance in language testing and assessment[J]. Lan-guage Testing, 2021(2): 189-218.
15. Kunnan A J, Jang E E. Diagnosing feedback in language assessment [A].In Long M H & Doughty C J. Handbook of Language Teaching [C]. New York: Wiley-Blackwell,2011.610-627.
16. Lee Y W, Sawaki Y. Cognitive diagnosis approaches to language assessment: An overview[J]. Language Assessment Quarterly,2009,6(3):172-189.
17. Wu Zunmin, Student Academic Achievement Evaluation (SAAE): A diagnostic assessment of Middle School English learning based on the National English Curriculum Standards,2011, Foreign Language Learning Theory and Practice,2011(04):15-22.
18. Sun Hang. (2019). Investigating students' cognitive processes in a diagnostic EFL reading test, Foreign Language Education in China,2019 2(04):25-32.
19. Leighton, J. P., M. J. Gierl. 2007. Verbal reports as data for cognitive diagnostic assessment [A]. In J. P. Leighton & M. J. Gierl (eds.). Cognitive Diagnostic Assessment for Education: Theory and Applications [C]. Cambridge: Cambridge University Press. 146-172.
20. Chen, J., J. de la Torre & Z. Zhang. 2013. Relative and absolute fit evaluation in cognitive diagnosis modeling [J]. Journal of Educational Measurement 50: 123-140.
21. Meng Yaru. Developing A Model of Cognitive Diagnostic Assessment for College EFL Listening[D], Shanghai: Shanghai International Studies University,2013.
22. DU Wenbo, MA Xiaomei.CDA-based diagnostic model construction for EFL reading tests[J], Foreign Language Teaching and Research,2018,74-88.