Language Testing in Asia

RESEARCH

Open Access

CrossMark

Developing reading proficiency scales for EFL learners in China

Yongqiang Zeng[1*] and Tingting Fan[2]

* Correspondence:
zengyq@gdufs.edu.cn
[1]Guangdong Teachers' College of Foreign Languages and Arts, 463 Shougouling Road, Tianhe District, Guangzhou, Guangdong, People's Republic of China
Full list of author information is available at the end of the article

Abstract

Background: Since its publication in 2001, the Common European Framework of Reference for Languages has exerted tremendous influence on language learning, teaching, and assessment in China. However, problems were also identified during its application process, thus calling for the localization of the framework. In 2014, the China Standards of English (CSE) project was launched, within which reading scales played a vital part.

Methods: Following the practice of CEFR, the CSE reading scales employed a similar methodology with revisions. In the initial phase, the theoretical framework was developed and operationalized. Then, a database of reading descriptors was built from both literature and empirical investigations. Finally, qualitative and quantitative measures were implemented to investigate and validate the classification of the descriptors into different categories at distinct levels.

Results: The paper reports the results of the first two research questions. Based on the synthesis of literature, reading proficiency was defined and four parameters were elicited to describe reading abilities of Chinese EFL learners. A total of 14,467 descriptors were collected, among which 13,893 were deleted after empirical investigations and iterative revisions, thus leaving 574 descriptors in the database.

Conclusions: Taking 3 years in construction and validation, the reading scales, together with other scales on ability and knowledge within the CSE, will exert tremendous influence on English learning, teaching, and assessment in the nation. However, problems are also expected, which require follow-up research and further revisions based on the feedback of the current version.

Keywords: China Standards of English, Reading scales, Cognitive abilities, Reading strategies

Background

The 1990s witnessed a mushroom increase in the design and development of language proficiency scales, among which are the construction and application of Interagency Language Roundtable (ILR), the International Second Language Proficiency Ratings (ISLPR), the Eurocentres Scale of Language Proficiency, the Canadian Language Benchmarks (CLB), the Common European Framework of Reference for Languages (CEFR), and the Global Scale of English (under development), to name just a few (see also Fan and Jin 2010; Han 2006; Wang 2012 for comprehensive reviews).

In mainland China, no language proficiency scales exist nationwide, except for several documents that serve as guiding principles for English education. The most

important are the *New English Curriculum for Chinese Primary Schools and Junior/ \Senior Middle Schools* (Ministry of Education, P. R. China 2012), the College English Curriculum Requirement (Ministry of Education, P. R. China 2007), and the Teaching Syllabus for English Majors in China (ELT Advisory Board under the Ministry of Education 2000).

Considering the demands of English development in the nation and acting as a positive response to the *Implementation Opinions in Deepening Reform on Examination and Recruitment System* (Political Bureau of the CPC Central Committee 2014) issued by the State Council on September 3, 2014, the China Standards of English (CSE) project was launched with general goals of benefitting various stakeholders in language learning, teaching, and testing (see Jin et al. 2016, this volume, for detailed information), although challenges were expected as those encountered in a decade's endeavor of the CEFR which has set a fine example for scale development and exerted tremendous influence on language learning, teaching, and assessment in China (Yang and Gui 2007; Fang et al. 2008; Cai 2012, 2013).

In the revised version of the College English Curriculum Requirement (Ministry of Education, P. R. China 2007), the "action-oriented approach" of CEFR was introduced and college English learners were viewed as social agents who have tasks to accomplish in their future studies, careers, and social interactions. Accordingly, the "Can-do" statements were applied in Chinese to describe undergraduates' EFL abilities with positive, clear, and definite expressions. In addition, self-assessment and peer-assessment scales were added to help learners profile their English skills, regulate their learning behaviors, and track their learning progress (Cen and Zou 2011). During the modification, however, problems emerged because CEFR's description of language proficiency from reception, production, and interaction and its contextualization of activities within public, personal, occupational, and educational domains were beyond the current practice of EFL learning and teaching in China and failed to cover important aspects that would influence the educational quality of the nation (Cen and Zou 2011). Similar concerns were also expressed by Zou et al. (2015) and Chen (2016).

Given the above problems, the CSE project was launched to describe the actual situation in the nation with essence drawn from the CEFR. In the CSE, language proficiency is considered as the ability to use it in a given set of circumstances (thus the "use-oriented" scale system) and is described on the basis of the following three perspectives: comprehension (consisting of listening comprehension and reading comprehension), production (comprising oral production and written production), and mediation (namely, translating and interpreting). In addition to the main constituents are parameters that influence different abilities, including grammatical knowledge, sociolinguistic knowledge, and pragmatic knowledge (see He and Chen 2016 in this volume).

As one of the main components of the CSE system, reading scales have been explored from the following three aspects:

(1) What is the theoretical basis for developing reading scales?
(2) What are the parameters for describing reading proficiency?
(3) To what extent can the reading descriptors be scaled and validated?[1]

Methods

Participants

Four groups of participants took part in the development of CSE reading scales:

(1) Core members (including the authors of this paper) that acted as designers, trainers, and data analysts for the whole reading project. On the one hand, they drew detailed blueprints and schedules during every phase of the project. On the other hand, they collected source literature, prepared materials, presided over the training session, and analyzed such qualitative and quantitative data as interviews with experts and descriptors compiled by teachers of various learning stages.

(2) Working group that was responsible for the compilation of reading descriptors. Particularly, 12 teachers from primary school to college were chosen as group leaders that were required to invite more teachers in the writing of descriptors. All teachers were recruited on the basis of their teaching experience and familiarity with learners' reading abilities at corresponding levels they were required to describe. Table 1 shows their background information and responsibilities in the writing task.

(3) Expert team that provided suggestions to the constructs of reading scales, parameters to describe reading proficiency, and categorization of the descriptors into distinct levels. They also participated, together with 12 teachers in the working group, in the interview sessions where questions concerning the feasibility of the framework and the readability of the descriptors were asked and recorded. All of the experts are professors in language testing or second language acquisition.

(4) Secretary team that consisted of eight postgraduates majored in language testing, responsible for inputting and arranging reading descriptors of multiple sources into an online platform where source and title of literature, access to the descriptor, original, translated and modified versions of the descriptor, level of the original descriptor, and name of the participant were required to be recorded. Averaged at 24 years old, all of them have learned English as a foreign language for over 11 years and received professional training on linguistics for 5 to 7 years (Table 2).

To ensure the quality of the descriptors for the reading scales, ideas were communicated effectively through emails and workshops within and across the four teams. Problems identified during the process were also exchanged with leaders of other groups (i.e., the listening group).

Materials

Materials analyzed in the present study include:

(1) Language proficiency scales that depict levels of language learning attainment at different stages (i.e., the Foreign Service Institute Scale, the Canadian Language Benchmarks, the Framework of the Association of Language Testers in Europe, the Common European Framework for Languages: Learning, Teaching, Assessment, etc.).

(2) Curriculum standards, for they reflect educational requirements of all learning stages. Examples are the New English Curriculum Standards for Chinese Primary Schools and Junior/Senior Middle Schools, the English Language Development

Table 1 Background information of the working group

Name	Gender	Age	Education background	Experience	No. of group members	CSE level for description	Target for description
Gao	Female	52	Postgraduate	Principle in one of the renowned primary schools in South China; rich experience in primary school English teaching	4	A1, A2	Primary school students
Wu	Female	48	Postgraduate	Principle in one of the average-level primary schools in South China; rich experience in primary school English teaching	3	A1, A2	Primary school students
Feng	Female	36	Postgraduate	Researcher in the Guangdong Education Bureau; rich experience in secondary school English teaching and training	22	A3	Secondary school students
Li	Male	26	Postgraduate	Outstanding young teacher in a province-level key school; rich experience in secondary school English teaching	9	A3	Secondary school students
Zhen	Male	56	Postgraduate	Researcher in the Guangdong Education Bureau; rich experience in high school English teaching and training	5	B1	High school students
Zhang (Q)	Female	42	Postgraduate	Researcher in the Zhuhai Education Bureau; rich experience in high school English teaching and training	5	B1	High school students
Pan	Male	39	Postgraduate	Teacher in one of the higher vocational colleges in South China; rich experience in higher vocational teaching and training	10	B2, B3	Undergraduates of non-English majors (including higher vocational college students); first- and second-year English major students
Zeng	Female	34	Postgraduate	Teacher in one of the higher vocational colleges in South China; rich experience in higher vocational teaching and training	10	B2, B3	Undergraduates of non-English majors (including higher vocational college students); first- and second-year English major students
Zhang (X)	Female	30	Doctoral candidate	Lecturer in one of the renowned universities in East China; a member of the CSE project	2	B3, C1	Third- and fourth-year undergraduates of non-English majors; undergraduates of English majors;
Xie	Female	36	Doctoral candidate	Lecturer in one of the renowned universities in South China; a member of the CSE project	4	C1,C2	Third- and fourth-year English major students; postgraduates of English majors; diplomatic personnel
Zhang (L)	Female	41	Doctor	Associate professor in one of the average-level universities in East	4	C2, C3	Postgraduates of English majors;

Table 1 Background information of the working group *(Continued)*

				China; a member in the CSE project			diplomatic personnel; senior translators
Cheng	Female	35	Doctor	Lecturer in one of the average-level universities in South China; rich experience in ESP and graduate English teaching	4	C2, C3	Postgraduates of English majors; diplomatic personnel; senior translators

(ELD) Standards, and the English Language Arts (ELA) Standards (including their implementation plans in the US states). Content-based as the ELA Standards are, they are involved in the present study for their emphases on critical-thinking and analytical skills, goals that are also set by the multi-functional CSE project.

(3) Test syllabi that provide useful information for language learning and teaching, such as syllabi for CET-4 and CET-6.

Table 3 presents examples of source materials.

Procedures

This study began with the survey of literature on reading theories and empirical studies of L2 reading ability, on the basis of which the theoretical model for developing reading scales was built, discussed, and refined. The final framework was operationalized into distinct parameters that were effective in describing EFL learners' reading ability.

A meta-study was then conducted to analyze reading proficiency scales and relevant literature. In this study, reading descriptors from existing proficiency scales, curriculum standards, and test syllabi at all the learning stages in China and out of the nation were assembled and classified (see the "Materials" section for detailed information). The inclusion and exclusion of the materials are based on rigorous literature review and expert judges. For instance, the ELA Standards of 19 states were selected for detailed analyses according to the studies by Wixson et al. (2003), Finn et al. (2006), and Wang (2012).

Meanwhile, 12 teachers from primary school to college were invited to write descriptors according to their teaching experience and their understanding of students' reading abilities. They were group leaders that were required to invite more teachers in the

Table 2 Participant information of the four groups

Group	Status	Participants		
		Gender	Number	Total
Core team	Professor; doctoral candidate	Female	4	5
		Male	1	
Working team	Teacher (from primary school to college)	Female	55	94
		Male	39	
Expert team	Professor	Female	3	8
		Male	5	
Secretary team	Postgraduate	Female	7	8
		Male	1	

Table 3 An example database for meta-analysis

Language	Title
Language proficiency scales	
English	The Foreign Service Institute Scale (FSI)
	The Interagency Language Roundtable (ILR)
	The American Council on the Teaching of Foreign Languages Guidelines (ACTFL)
	International Second Language Proficiency Ratings (ISLPR)
	Canadian Language Benchmarks (CLB)
	Framework of the Association of Language Testers in Europe (ALTE)
	The Common European Framework for Languages: Learning, Teaching, Assessment (CEFR)
	Finnish Scale of Language Proficiency (FSLP)
	Global Scale of English (GSE)
	The EIKEN Can-Do List (EIKEN)
Curriculum standards	
English	The Common Core State Standards (CCSS)
Chinese	English Language Development Standards (ELD)
	WIDA ELD Standards (ELD_WIDA)
	California ELD Standards (ELD_C)
	English Language Development Wiki (ELD_W)
	Core Academic Skills for Educators (CASE)
	Alabama English Language Arts Standard (AELA)
	Minnesota Academic Standards (MAS)
	Washington State's Essential Academic Learning Requirements (WSEALR)
	New English Curriculum for Chinese Primary Schools and Junior/Senior Middle Schools: 2011 Edition (NEC)
	Teaching Syllabus for English Majors in China (TSEM)
	College English Curriculum Requirement (CECR)
	Curriculum for Business English Correspondence (CBEC)
	Curriculum for Comprehensive English (CCE)
	Curriculum for English Writing (CEW)
	Curriculum for Non-English Major Postgraduates (CNMP)
Test syllabi	
English	YLE Handbook for Teachers (YLE)
Chinese	FCE Handbook for Teachers (FCE)
	KET Handbook for Teachers (KET)
	PET Handbook for Teachers (PET)
	BEC Handbook for Teachers (BEC)
	CAE Handbook for Teachers (CAE)
	CPE Handbook for Teachers (CPE)
	General Certificate of Secondary Education Syllabus (GCSES)
	Advanced GCE in English Language Specification (AGCE)
	Advanced Subsidiary GCE in English Language Specification (ASGCE)
	TOEIC Can-Do Guide Executive Summary (TOEIC)
	Syllabus for College English Test—Band Four (CET-4)
	Syllabus for College English Test—Band Six (CET-6)
	Syllabus for Test for English Majors—Band Four (TEM-4)
	Syllabus for Test for English Majors—Band 8 (TEM-8)
	Standards for Public English Test System (PETS)

writing task (94 in total). Before the assignment of tasks, the participants underwent a training session, which included the following steps:

(1) The voluntariness was acknowledged and the purpose of the study was conveyed, followed by an elaborate explanation of the theoretical framework, principles, and preliminary parameters in describing reading.

(2) Ten sample descriptors were distributed to the participants to familiarize themselves with the structure of descriptors. Questions about the framework and formula were raised and solved.

(3) Participants were required to write ten descriptors independently, followed by a discussion session where problems were identified and discussed.

At the end of the training, tasks were distributed. All descriptors were submitted by the end of July 2015.

The database built from the literature and teachers' writing was then subject to quantitative analyses during which all of the descriptors were coded and analyzed according to the parameters set forth in the "Parameters for Describing Reading Proficiency" section, as well as qualitative investigations when the descriptors were refined after a series of focus-group interviews with experts, teachers, students, and English learners of other working fields (see the "Refinement of Parameters" section). In the interviews, the readability, feasibility, and classification of descriptors as well as suggestions for further improvement were discussed and digitally recorded. All the transcribed data were cross-examined by authors of this article and emailed to the experts and teachers individually for authenticity check.

At the following stage, qualitative and quantitative measures have been implemented to investigate and validate the classification of descriptors into different categories at distinct levels with questionnaires containing both parallel and vertical anchor items scattered among teachers and students over the nation. In order to estimate and adjust for rater severity in calibrating the questionnaire results onto the scale, the multi-faceted Rasch model will be used as operationalized in the FACETS program. Other procedures will also be applied to examine the usefulness and effectiveness of the scales and explore the expansion of such use-based scales to diagnosis and other orientations, in which some initial results have been obtained within the group (i.e., T. Fan: Developing and validating diagnostic reading proficiency scales for Chinese EFL learners in high school, unpublished). Below is the flowchart that illustrates the progress of the project (Fig. 1).
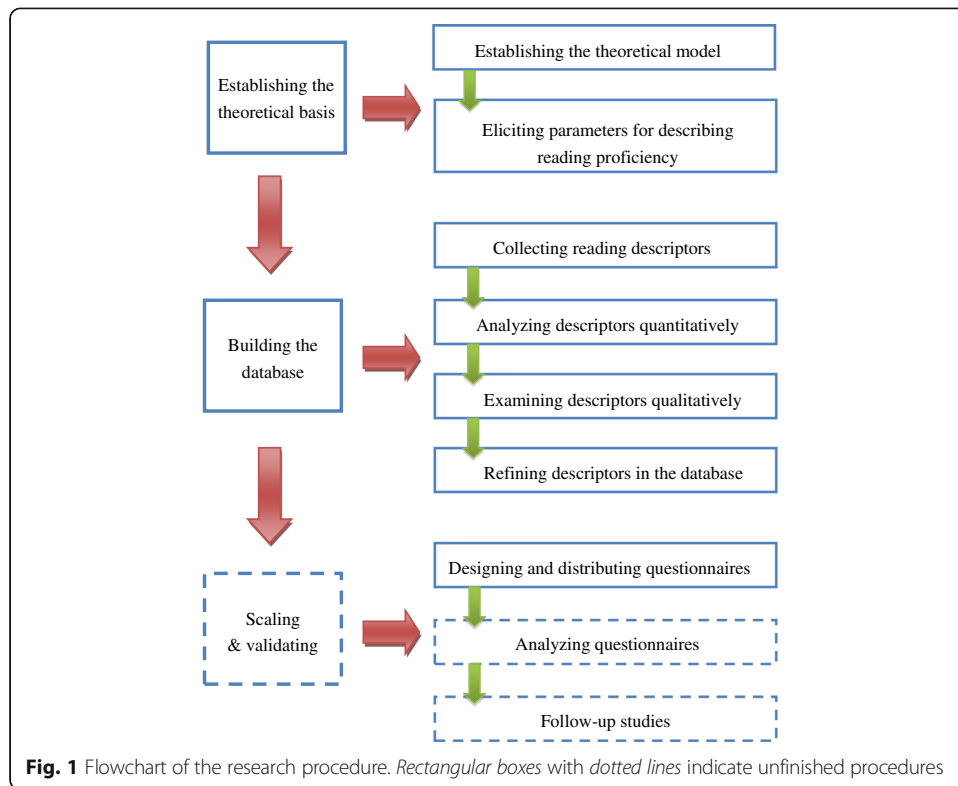
Results

Definition of reading comprehension

Reading comprehension has been explored by researchers over the years, during which diversified definitions have been provided by researchers and various models and theories have been proposed and entwined with general theories that influence education and psychology.

Alderson (2000) regarded reading as a complex process affected by such variables as reader (including knowledge, skills, abilities, motivation, affect, and other characteristics of the reader) and text (for example, topic and content, type and genre, text organization), while Koda took reading as the extraction and integration of information from the text on the part of readers and as combination of the newly acquired information with "what is already known" (2005, p. 4). Cognitive psychologists viewed reading as the internal process that requires readers to build mental model from the text and from information and visual clues outside of the text (Bernhardt 1991; Just and Carpenter 1987; Garnham and Oakhill 1996; Grabe 2009; Kintsch 2004; Zwaan and Radvansky 1998; Zwaan and Rapp 2006).

Based on the synthesis of literature, the present study defines reading as language users' or learners' disposal of their cognitive processes, comprehension strategies, and knowledge to construct meaning from written materials in various contexts and under various conditions. Cognitive ability is composed of a broad range of lower-level abilities and higher-level components (i.e., recognizing, understanding and analyzing, evaluating, and criticizing) and is finally grouped into six functions that each descriptor

Fig. 1 Flowchart of the research procedure. *Rectangular boxes* with *dotted lines* indicate unfinished procedures

is intended to convey. Comprehension strategies involve planning, executing, evaluating, and repairing, while knowledge includes both language knowledge (i.e., grammatical knowledge) and non-language knowledge (i.e., socio-linguistic and pragmatic knowledge) (National Education Examinations Authority: Workbook on China Standards of English, unpublished).[2] Figure 2 provides an illustration to the framework.

Fig. 2 Framework of the CSE reading scales
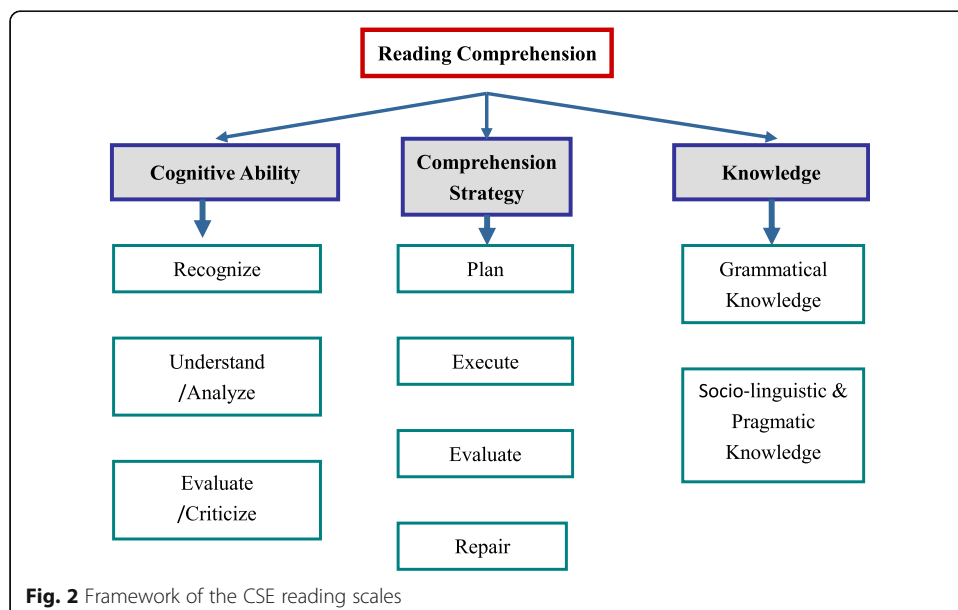
In this framework, recognizing refers to the recall and recognition of information, ideas, and principles in the approximate form in which language learners have acquired. Understanding is related to the translation and interpretation of information on the basis of prior learning. Analyzing implies the distinction, classification, and integration of assumptions, hypothesis, evidence, or structure of a statement. Evaluating and criticizing are the appreciation, assessment, or critique of the reading material based on specific standards and criteria. Strategies include a variety of measures that readers take during the failure of comprehension.

Three reasons account for the above definition and categorization of reading comprehension abilities: (1) the definition conforms to the theoretical conception of the CSE project which takes language proficiency as dynamic activities rather than abstract rule systems, and is consistent with the framework of other scales such as listening comprehension to build a systematic whole of the project (National Education Examinations Authority: Workbook on China Standards of English, unpublished). (2) It covers a comprehensive range of cognitive abilities, comprehension strategies, and language knowledge, subsuming both lower-level processes and higher-level skills explored by various models of reading comprehension. The multi-layered nature of the definition provides the feasibility for the current attempt in developing comprehensive reading scales and the possibility for future refinement of the scales in the light of different purposes (for diagnosis, for example). (3) The framework has been applied successfully in scale development in previous studies. For example, Wang (2012) divided language comprehension ability into four macro cognitive skills—recognizing and recalling, understanding and summarizing, analyzing and inferring, and critiquing and appreciating—on the basis of a synthesis of literature on cognitive abilities, reading, and listening comprehension.

Parameters for describing reading proficiency

At this stage, principles were set following the practice of previous scales (i.e., CEFR and GSE). Two dimensions were identified to depict the development of reading proficiency—quantity and quality, with the former stipulating how many things language learners can do while the latter describes the accuracy, fluency, or confidence that learners can do them. Quality can be further divided into intrinsic quality, that is, the criteria of the performance, and extrinsic quality, or conditions under which the performance is completed (de Jong 2015). These two dimensions are merged into descriptors of reading proficiency where the verb and goal (object of the verb) express the particular task that a reader is required to complete and the modifier and text parameters specify the difficulty of materials that the reader has to process, or the conditions under which the task is performed (see Table 4 for illustration).

Refinement of parameters

By the end of July 2015, a total of 14,467 descriptors were collected, among which 1398 were summarized from literature by both core and secretary members and 13,069 were compiled by teachers. All of the descriptors thus collected underwent iterative sessions of close examination by experts on reading comprehension, during which all descriptors were coded and analyzed according to the verb, goal, modifier, topic, and text type parameters set previously. An example of the coding is:

Table 4 Parameters of reading descriptor

Parameter	Quantity/performance		Quality/criterion	
	Verb	Goal	Modifier	Topic and text type
Structure	Can (understand)	sth.	In an adj.	On a certain topic/material
Example (cognitive ability)	Can recognize	Specific information	In a simple	Story about personal experience
Example (comprehension strategy)	Can analyze	The author's attitude	By punctuations in	Stories about life experience and perception

Can <appreciate> [aesthetic values] in a (literary work) with /rich connotations/.

where < > signifies the verb of the descriptor, [] indicates the object or goal of the predicate, () represents the particular text or specific topic the reader is required to comprehend, and // designates the language, organization, and other modifiers of the reading material.

The coding results revealed a confused picture. For the first parameter, a total of 52 verbs were identified to describe cognitive abilities and comprehensive strategies employed in reading, followed by a massive number of goal expressions that varied from detailed information to writer's attitudes and opinions, in order to accommodate to the particular predicate before them. For the parameter of modifier, the more complicated situation emerged because 265 phrases were used in defining the text that readers were able to comprehend. The phrases could be further classified into four sub-categories—language (difficulty, complexity, style, etc.), content and theme (complexity, abstractness, etc.), length of the text (short, medium, long), and method of reading (fast reading, careful reading, etc.). The same case occurred on the topic and text type, for a multitude of topics scattered in six types of texts and ranged from the abstract (i.e., argumentation, exposition) to the concrete (i.e., manual, comment on current affairs, business letter).

Afterwards, focus-group interviews were conducted among experts and teachers to investigate the practicality of the framework, the structure of descriptors, and the coherence of expressions. For the reading framework, the interviewees agreed that the cognitive pyramid in the theoretical framework provided a sound basis to distinguish reading abilities between levels. Take primary school students (i.e., readers at A1 and A2) for example. At the bottom of the cognitive pyramid, they are confined to "recognize" and "retrieve" textual information from simple texts, unable to analyze text structure or evaluate the writer's purpose. The results justify the feasibility of the theoretical framework in developing reading scales. However, confusion was also expressed on the expression of the four parameters (as consistent with coding results) and the structure of descriptors. Major concerns include:

(1) Expressions of the goal parameter should be constrained within reading proficiency, especially for descriptors on reading strategies, because some were depicted from the perspective of language learning rather than reading, as demonstrated by such phrases as "sharing information with peers" and "by referring to literature on related topics".

(2) The reading texts described by each level were not typical enough for language learners to comprehend. It would be sufficient for lower level readers to comprehend the overall idea of picture books and nursery rhymes. For those at higher levels, particularly, English major students at college, they are required to process texts of various types and genres, including news reports, original novels, and essays.

(3) A total of six text types were categorized and defined in the CSE reading system after discussion. The definition of each text type is presented below.

- Narrative text: It recounts a sequence of events (real and imagined included). Typical examples are stories, fictions, poetries, and so on and so forth.
- Descriptive text: Often embedded in other types of writing, descriptions, as the name suggests, are used to describe a place or to create a particular mood and atmosphere so that readers can visualize vivid pictures of the character, the place or the object.
- Expository text: This type of text explains the shape, structure, category, and relation or function of a particular object or illustrates the definitions and features of certain theories. Examples are news reports, technical instructions, and academic lectures.
- Argumentative text: It expresses the writer's viewpoint by presenting relative evidence to convince the audience of a particular stance. The drafts of speech and debate fall into this category.
- Instructional text: Instructions express operational and instrumental functions in providing directions for the reader. Manuals, recipes, and announcements are the texts of this type.
- Social text: It is utilized to establish, maintain, or change the interpersonal relationships between the writer and the intended reader. It usually appears in oral interactions. In written materials, mails and emails are appropriate examples.

(4) Consensus was reached to delete text parameters in strategy description, for strategies are always activated in specific reading tasks and the involvement of text information seems to be redundant. Therefore, such descriptor as "Can understand the relationship between sentences or paragraphs by cohesive words and phrases in argumentative texts of medium difficulty" was changed into "Can understand the relationship between sentences or paragraphs by cohesive words and phrases."

On the basis of coding results and interview data, decisions were made to refine parameters in describing reading proficiency. For cognitive abilities, five parameters were selected to describe reading—verb, goal, text type, topic, and language, among which the first three are essential to every descriptor while the last two are added if the compulsory elements fail to distinguish reading abilities between levels. For beginners, the compulsory elements would be sufficient to describe reading abilities. With the increase of ability levels, however, compulsory elements would be insufficient and other parameters must be added in the description. In addition, consensus has been reached to delete text parameters in strategy description. Tables 5 and 6 display the revised versions of ability and strategy description.

In the light of the revised parameters, all descriptors collected beforehand underwent two rounds of close examinations. In the first round, eight experts and teachers

Table 5 Revised parameters of reading descriptor: cognitive ability

CSE level	Function	Ability	Quantity/performance		Quality/criterion		
			Verb	Goal	Text	Topic	Language
A1	Narration	Recognize	Can recognize	Rhyming words	In children's poems		
C2	Argumentation	Understand/ analyze	Can understand	The organization	Of academic monographs	In professional fields	
A3	Narration	Recognize	Can retrieve	The main plot	In novels		Of the simplified version
B2	Argumentation	Understand/ analyze	Can analyze	The implicit attitudes of the writer	In comments	On current affairs	
B2	Interactiveness	Evaluate/ criticize	Can evaluate	Expression of personal feelings	In practical writing like resume and cover letter		

selected from the expert and working teams were invited to revise the descriptors and delete those that lacked readability and practicability or that did not conform to the structure of descriptors, giving rise to 4884 descriptors in the database. The second round was implemented by the group leader and core members to re-examine the descriptors, leaving a total of 574 descriptors in the database. All of the descriptors were classified into two broad categories of cognitive abilities (which were further divided according to the function that each descriptor conveyed) and reading strategies and into hypothetically nine CSE levels (ranging from A1 to C3) defined on the basis of learning stages in the nation for the convenience of data collection. However, these levels are subject to empirical justification and adjustment in the scaling phase of the project.

Discussion

Since the breakthrough made by the FSI scale that filled the void of systematic descriptions for language use, proficiency scales and documents undertaking similar functions have been issued in Northern America, Europe, Australia, and other nations, among which the most renowned is the CEFR that has left an indelible mark on language learning, teaching, and assessment in China. However, the philosophies advocated by the CEFR are beyond the current practice of Chinese EFL learners and teachers, calling for the localization of the framework and thus giving rise to the CSE project that was launched in 2014 with the aims of improving English proficiency of individual learners and educational quality of the nation, providing a theoretical framework for English teachers to refer to, and aligning the teaching and testing practice throughout all the stages from pre-schooling to higher education. As major components of the CSE,

Table 6 Revised parameters of reading descriptor: comprehension strategy

CSE level	Ability	Quantity/performance		Quality/criterion
		Verb	Goal	Instrument
A3	Execute	Can understand	The written material	By group discussion
B2	Evaluate	Can evaluate	The author's opinions	By argument and supporting evidence
A2	Repair	Can understand	The written material	With the help of explanatory notes

reading scales are to be issued in 2017 as a result of collective endeavor. Compared with previous scales, the CES reading scales are improved in the following aspects.

The first set of improvements is concerned with the "theoretical validity" (Weir 2005). Nearly half of the previous scales lack sound theoretical bases. In ACTFL Guidelines (ACTFL 2012), for example, it is openly stated that "the Guidelines are not based on any particular theory, pedagogical method, or educational curriculum." Even though another half assumes their foundations with communicative competence and language development theories, the ground is still on general language ability theories applicable for all the skills of listening, speaking, reading, and writing. The present project makes a tentative attempt in the development of reading scales with a synthesis of literature on cognitive abilities and reading theories. The framework provides the possibility to explore reading scale development at finer levels than previous endeavors.

Secondly, the reading scales abandon the native-speaker standard adopted by some previous scales like the FSI family. The Handbook of ACTFL Guidelines explains this "educated native speaker norm" when stating "educated refers not to holding a degree or diploma," but to "using the speech or language generally associated with educated speakers in the native country." Preference is thus given for the language variety of a particular social group within the second-language culture. In this view of language use, these foreign language scaling systems are thus clearly elitist, not adaptable for school programs. Bachman (1990) also considered this norm inadequate because native speakers showed considerable variation in ability, particularly with regard to comprehension abilities that required cognitive resources from working and long-term memory. "Given the number of varieties, dialects, and registers that exist in virtually every language, we must be extremely cautious in attempting to treat even 'native speakers' as a homogeneous group" (p. 39). In line with Bachman's conceptions, reading proficiency in the CSE is defined in terms of abilities instead of actual performance of native speakers, thus offering the potential for accurate descriptions.

Finally, the descriptors are grouped by the function that each descriptor is intended to convey, thus providing a comprehensive description of EFL learners' reading ability, in contrast with previous scales that concentrate on particular domains. In the ALTE framework, for example, reading ability is specified and described in three contexts separately—social and travel contexts, the workplace, and studying, providing convenience for users to make reference to descriptors in particular areas. The same occurs in the CEFR where reading comprehension is described through visual reception activities that readers "receive and process as input written texts produced by one or more writers" (Council of Europe 2001, p. 68). Therefore, in addition to the overall scale depicting language learners' abilities to read for gist, for information, for detailed understanding, and for implications when they are dealing with texts of different topics at various difficulty levels, four subscales are developed to portray their performance hierarchically on four typical activities—reading correspondence, reading for orientation, reading for information and argument, and reading instructions. Such descriptions are insufficient for language learners in China (see the introduction part for detailed illustration).

Conclusions

Starting from the synthesis of literature on cognitive abilities and reading theories, reading proficiency was defined, in the CSE, as language users' or learners' disposal of their

cognitive processes, comprehension strategies, and knowledge to construct meaning from written materials in various contexts and under various conditions, from which four parameters were elicited to describe reading abilities of Chinese EFL learners—verb, goal, modifier, topic, and text type. Then, the database with 14,467 descriptors was built and 574 descriptors remained after empirical investigations and iterative revisions.

Then, qualitative (i.e., interviews) and quantitative (i.e., questionnaires) measures have been implemented across the nation to investigate and validate the classification of the descriptors into different categories at distinct levels (to be completed). All the data will be coded systematically and analyzed scientifically.

Starting from 2014, the reading scales, together with other scales on ability and knowledge within the CSE, will exert tremendous influence on English learning, teaching, and assessment in the nation. However, problems are also expected, which require follow-up research and further revisions based on the feedback of the current version.

Endnotes

[1]Since the project has not been finished, this paper reports the results of the first two questions. Nevertheless, the third research question remains here as it plays a vital role in the scale development.

[2]Although knowledge is one of the three major components of the CSE reading framework, it is investigated and described by the organizational knowledge group according to the assignment of the project.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]Guangdong Teachers' College of Foreign Languages and Arts, 463 Shougouling Road, Tianhe District, Guangzhou, Guangdong, People's Republic of China. [2]Nanjing University of Aeronautics and Astronautics, 29 Jiangjudadao Road, Jiangning District, Nanjing, People's Republic of China.

References
Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
ACTFL. (2012). *ACTFL proficiency guidelines*. http://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012. Accessed 05 Oct 2014.
Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
Bernhardt, E. B. (1991). A psycholinguistic perspective on second language literacy. *AILA Review, 8*, 31–44.
Cai, J. (2012). The impact of CEFR on foreign language teaching in China. *China University Teaching, 6*, 6–10.
Cai, J. (2013). Adjustment of educational assessment of EFL at tertiary level against the background of internationalizaiton of higher education. *Comput-Assist Foreign Lang Educ, 149*, 3–8.

Cen, H., & Zou, W. (2011). An impact study of the CEFR on Chinese college English education. *Foreign Lang China, 8*(4), 31–38.

Chen, X. (2016). CEFR: a guideline on business English teaching and testing. *Foreign Lang Res, 3*, 133–136.

Council of Europe. (2001). *Common European framework of reference for languages: learning, teaching, assessment*. http://www.coe.int/t/dg4/linguistic/source/Framework_EN.pdf. Accessed 05 Oct 2014.

de Jong, J. (2015). *Introduction to understanding and using the common European framework*. Beijing: Report presented at the meeting of the CSE project.

ELT Advisory Board under the Ministry of Education. (2000). *Teaching syllabus for English majors in China. Beijing: foreign language teaching and reach press*. Shanghai: Shanghai Foreign Language Education Press.

Fan, J., & Jin, Y. (2010). A study of language testing standards: review, reflection and inspiration. *Foreign Lang World, 1*, 82–91.

Fang, X., Yang, H., & Zhu, Z. (2008). Creating a unified scale of language ability in China. *Mod Foreign Lang, 31*(4), 380–387.

Finn, C. E., Julian, L., & Petrilli, M. J. (2006). *The state of state standards*. https://edexcellence.net/publications/soss2006.html. Accessed 22 Oct 2014.

Garnham, A., & Oakhill, J. (1996). The mental models theory of language comprehension. In B. K. Britton & A. C. Graesser (Eds.), *Models of understanding text* (pp. 313–339). Mahwah: Erlbaum.

Grabe, W. (2009). *Reading in a second language: moving from theory to practice*. Cambridge: Cambridge University Press.

Han, B. (2006). A review of current language proficiency scales. *Foreign Language Teaching and Research, 38*(6), 443–450.

He, L., & Chen, D. (2016). Developing a common listening ability scales for Chinese learners of English. Language Testing in Asia, this volume

Jin, Y., Wu, Z., Alderson, C., & Song, W. (2016). Developing a framework of reference for English language education in China: challenges at macro- and micro-political levels. Language Testing in Asia, this volume.

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Old Tappan: Allyn & Bacon.

Kintsch, W. (2004). The construction-integration model of text comprehension and its implications for instruction. *Theoretical Models and Processes of Reading, 5*, 1270–1328.

Koda, K. (2005). *Insights into second language reading: a cross-linguistic approach*. Cambridge: Cambridge University Press.

Ministry of Education, P. R. China. (2007). *College English curriculum requirement*. Beijing: Higher Education Press.

Ministry of Education, P. R. China. (2012). *New English curriculum for Chinese primary schools and junior/senior middle schools* (2011th ed.). Beijing: Beijing Normal University Publishing Group.

Political Bureau of the CPC Central Committee. (2014). *Implementation opinions in deepening reform on examination and recruitment system*. http://www.gov.cn/zhengce/content/2014-09/04/content_9065.htm. Accessed 08 Oct 2015.

Wang, S. (2012). *Developing and validating descriptors of language comprehension ability for Chinese learners of English*. Beijing: Intellectual Property Publishing House.

Weir, C. J. (2005). Limitations of the common European framework for developing comparable examinations and tests. *Lang Test, 22*(3), 281–300.

Wixson, K. K., Dutro, E., & Athan, R. G. (2003). The challenge of developing content standards. *Rev Res Educ, 27*(1), 69–107.

Yang, H., & Gui, S. (2007). On establishing a unified Asian level framework of English language proficiency. *Foreign Lang China, 4*(2), 34–37.

Zou, S., Zhang, W., & Kong, J. (2015). On the research status and application prospects of CEFR in China. *Foreign Lang in China, 12*(3), 24–31.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychol Bull, 123*(2), 162.

Zwaan, R. A., & Rapp, D. N. (2006). Discourse comprehension. *Handbook of Psycholinguistics, 2*, 725–764.