Check for updates

Verbal reports in the reading processes of language learners: A methodological review

Patriann Smith[1],[*] (iD), Deoksoon Kim[2], Oksana Vorobel[3] and James R. King[4]

[1]*Department of Teaching and Learning, College of Education, University of South Florida, Tampa, FL, USA,* [2]*Teacher Education, Special Education, Curriculum & Instruction, Lynch School of Education and Human Development, Boston College, Chestnut Hill, MA, USA,* [3]*Borough of Manhattan Community College, City University of New York, New York, NY, USA,* [4]*University of South Florida, Tampa, FL, USA*

This methodological review highlights the trends in empirical studies where a methodological construct (i.e. verbal reports) intersects with content (i.e. literacy research). Specifically, we synthesise research on language learners' reading in which verbal reports were deployed as a methodological tool. Questioning the long-standing assumption that verbal report methods validated in first-language verbal report reading studies necessarily constitute a basis for validation of second-language verbal report reading studies, we consider the broader educational frameworks within which studies are embedded. In our synthesis of language learners' verbal report literacy research published between 2000 and 2015, we attend to the social, demographic and geographic realities characteristic of studies reviewed and of the participants involved. Our findings demonstrate the following: (a) tendency to report quantitative information regardless of the type of verbal reporting method and the component of reading explored; (b) predominance of independent concurrent methods that emphasised the reading *product*; (c) predominance of integrated verbal reports (i.e. concurrent and other forms of reporting) in sociocultural studies that reflected the reading *process*; (d) concerns about validity in studies premised on cognitivist models of verbal reports; (e) a tendency to use solely concurrent verbal reports in quantitative studies; (f) high reliance on integrated concurrent methods in qualitative studies; and (g) preponderance of qualitative-to-quantitative versus a qualitative-to-quantitative-to-qualitative verbal reporting paradigms across studies. Based on these findings, we make several recommendations to be considered when verbal reports are used to study language learners' reading processes.

English language learners (ELLs) are the fastest growing student population in countries such as the USA, the UK, and Canada (National Clearinghouse for English Language Acquisition [NCELA], 2011; European Commission, 2012; People for Education, 2013; National Center for Educational Statistics 2018). As a result, educators have paid increasing attention to the underlying literacy processes inherent in the multiple linguistic repertoires of this population (Grabe, 2009; Han & Anderson, 2009; Bernhardt, 2011). Both nationally and internationally, there is a search for resources and opportunities that may enhance language learners (LL's) literacy development (National Council of Teachers of English, 2008; Common Core State

*Corresponding author. Patriann Smith, Department of Teaching and Learning, College of Education, University of South Florida, 4202 E. Fowler Ave., Tampa, FL 33620, USA. Email: patriann. s.smith@gmail.com

Standards Initiative, 2010; TESOL, 2010; Risko *et al.*, 2017). Many researchers have explored reading comprehension and literacy practices of LLs (Fitzgerald, 1995; Snow *et al.*, 1998; NCTE, 2008, 2011). Despite these efforts, we still know little about the patterns derived from the use of certain prominent methodologies for studying LLs' literacy development.

One prominent methodology for examining language learners' (LLs') literacy development is verbal reports. Verbal reports constitute individuals' oral descriptions of the mental processes in which they are engaged (Ericsson & Simon, 1984). Historically, verbal protocols, as they are sometimes referred to, have been used widely to understand first-language (L1) students' literacy processes (Pressley & Afflerbach, 1995). In their use of such protocol analysis, studies primarily reflect a cognitivist approach, one in which reports are considered windows on thought processes as they occur (Bowles, 2010a,b). However, in the more contemporary literature, with the increased attempts to understand literacy practices of LL and with the definition of literacy expanding to reflect multiple literacies and how context affects literacy (Firth & Wagner, 1997; Leu *et al.*, 2004), studies have increasingly used verbal reports to understand the compensatory processing between L1 and L2 and information processing, that is, factors influencing reading that go beyond those inferred from the mind in which second-language (L2) reading is embedded (e.g. Gass & Mackey, 2007; Bernhardt, 2011).

Adopting verbal protocol methods can provide a window to examine social factors involved in ELLs' reading comprehension (Kim, 2011; Smith & King, 2013; Kim, 2018). We are beginning to learn more about how verbal reports may provide insight not only into LLs' cognition but also about how they mobilise social tools to facilitate literacy development (Bowles, 2010a; Cohen, 1983, 1987, 1996, 2013; Vorobel & Kim, 2017; Kim, 2018). The calls for greater attention to social processes involved in using verbal reports for second-language reading have resulted in studies such as those by Kim (2018) that point out the importance of integrating cognitive and sociocultural perspectives. Kim (2018) identifies two broad types of reading strategies used by ELLs: *higher order thinking strategies* (e.g. interferential thinking, intertextuality, comprehension monitoring, evaluating) and *social–contextual reading* strategies (e.g. connecting to oneself, connecting to the world, socio-affective strategies). She notes that beginning ELLs often favour social–contextual reading strategies while more advanced ELLs more often use higher order thinking strategies. All readers combine cognitive and sociocultural approaches, and it is most productive for research in second-language literacy to attend to both (Kim, 2018).

Though studies such as this highlight attention to additional factors involved in ELLs' reading processes, little has been synthesised about how studies have used verbal reports to explore LLs' reading processes in what is now considered the era of new literacies (Leu *et al.*, 2004). Moreover, previous reviews of L2 studies focused on such topics as cognitive reading processes (Fitzgerald, 1995), writing strategies (e.g. Matsuda *et al.*, 2003; Silva & Brice, 2004), general strategy use (e.g. Chamot, 2004), deployment of think-alouds (Bowles, 2010a), and reading programs (e.g. Cheung & Slavin, 2012). These reviews emphasise the cognitive reading processes in which second-language readers engage.

This methodological review therefore fills a gap in the literature by highlighting trends in original studies where verbal reports have been used in literacy research with the intention to recommend how they may best be used with LLs. Our review thus examines research on LLs' literacy published between 2000 and 2015 that used verbal reports as a methodological tool. We examine the educational, social, cognitive, demographic, economic, and geographic realities that bear upon the ways in which such students' literacy practices are examined as well as synthesise trends in the use of verbal report methodology in reading studies conducted with LLs. The following overarching questions guided this review:

1 How have verbal reports been used in empirical studies to explore LLs' reading processes?
2 What social, cognitive, geographic, and demographic patterns and concerns present themselves in empirical studies where verbal reports have been used to explore LLs' reading processes in various contexts, with multiple languages being used by different learners in different contexts?

Definitions of language learners

The term, English language learner (ELL), has been used to refer to 'active learner of the English language who may benefit from various types of language support programs' (NCTE, 2008, p. 2); as a bilingual who possesses emerging English proficiency (Bialystok, 2001); or as a 'non-native English speaker(s) learning English in school' (Peregoy & Boyle, 2008, p. 2). For the purpose of this review, we refer to ELLs as non-native English speakers 'learning English skills and knowledge' (Cheung & Slavin, 2012), who, regardless of the setting in which they learn language or the age at which a new language is learned, possess varying degrees of proficiency in one or more other languages. Similarly, we use the term *language learners* (LLs) to refer to students learning the skills and knowledge of a new language, and *second-language (L2) learners* to refer to students learning a second language where this is the dominant language in the social context. The terms, *third-* and *fourth-language learners* (L3, L4 respectively), represent learners similar to L2s. However, we make the distinction that L3 and L4 learners acquire multiple languages simultaneously, or continuously learn a third and/or fourth language in addition to possessing competence in two languages. *Foreign language* (FL) *learners* are taken to mean those students learning a new language in an environment where their L1 is the principal language spoken.

Variations of verbal reports

To facilitate our discussions of verbal reports, we identify various ways of implementing this methodological tool. Among these variations are concurrent (or non-metalinguistic), retrospective, introspective (or metalinguistic), and delayed retrospective types of verbal reports. *Concurrent (or non-metalinguistic) reports* are produced simultaneously with the academic task being completed, with a primary focus on the task and reporting ongoing thoughts as secondary. *Retrospective reports* occur after the task has been completed, with a delay between task completion and verbal reporting.

Introspective reports ask the participant to explain how s/he developed a particular con-current thought, while *delayed retrospective reports* occur after a specific amount of time has elapsed following the academic task (Ericsson & Simon, 1993).

Method

Literature search procedures

We first developed a list of journals for our literature search, based on Smith and Lafford's (2009) recommendations and our observations from reviews and meta-analyses of research in second language acquisition (SLA) (e.g. Fitzgerald, 1995; Norris & Ortega, 2006; Bowles, 2010a; Biber *et al.*, 2011). Based on these insights, 15 journals constituted the list: *Applied Linguistics, CALICO Journal, Canadian Modern Language Review, Computer Assisted Language Learning, Computers and Education, English Language Teaching Journal, International Review of Applied Linguistics, Language Learning, Language Learning and Technology, Multilingual Education, Reading Research Quarterly, ReCALL, Studies in Second Language Acquisition, System* and *TESOL Quarterly*. Of these journals, 12 reflected articles that met our inclusion criteria (discussed later), while three, namely *ReCALL, Multilingual Education* and *Computers and Education*—did not yield any studies that met the inclusion criteria set for the review. Beyond the journals selected for our search, we then conducted an independent search of the PsycInfo and Education Resources Information Center (ERIC) databases because they contain many journals in the field of education and applied linguistics and are frequently used in systematic reviews and meta-analyses (Lin *et al.*, 2013; Huber & Kuncel, 2016).

Second, we used the terms: (a) *verbal reports*, (b) *think-alouds*, (c) *reading*, (d) *verbal protocols*, (e) *reading process*, (f) *second language*, (g) *bilingual*, (h) *multilingual*, (i) *English as a Second Language (ESL)* and (j) *foreign language*, along with their combinations, to search the indexes of the selected journals and the databases back to 2000. The inclusion criteria required articles to (a) report an empirical study, (b) include a form of verbal report methodology that occurred in conjunction *the product of reading* with exploration of the reading task being studied, (c) focus on LLs' reading processes, (d) involve research conducted within the K-university levels and (e) have a publication date between 2000 and 2015—from the beginning of the century that coincided with the influential discussion on sociocultural theory by Firth and Wagner (1997, 1998)—and others up to the last complete year of publications before the preparation of this article. This discussion proposed a reconceptualisation of research in SLA to extend examinations of language use and acquisition beyond cognitivist perspectives that often reflect deficit orientations of language learners, towards orientations that foregrounded the social and contextual factors involved in this process (Firth & Wagner, 1997). The exclusion criteria required us to reject studies where verbal reports were: (a) not used in conjunction with a reading task or (b) geared toward an understanding of testing characteristics of participants, rather than reading processes. Using these inclusion and exclusion criteria, we identified 76 empirical studies that met our criteria (see Table 1 for a detailed overview of the studies). We included geographical regions beyond the USA.

Table 1. Summary of key features of original studies.

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Akpinar (2013)	N/A	$N = 40$ Ages: 18–22 Language(s): EFL, Turkish as L1 Language Level: intermediate Academic Level: undergraduate Department: English Language Teaching Location: Gazi University, Turkey	To investigate EFL learners' perceptions and practices when guessing the meaning of unfamiliar words during reading.	Mixed methods. A two-fold vocabulary strategy survey, reading task, a lexical inference test with concurrent verbal reports; Pearson correlation coefficients, coding, intercoder agreement.	The majority of EFL learners preferred looking up an unknown word in an English–Turkish dictionary and guessing the meaning from the context. There was an insignificant correlation between Turkish L2 learners' perceptions and actual practices for the contextual and intralingual knowledge sources.
Akyel and Erçetin (2009)	N/A	$N = 10$ Ages: 21–24 Language(s): ESL Language Level: advanced Academic Level: undergraduate Department: ELT Location: Turkey	To examine advanced L2 readers' processing strategies in reading hypermedia text.	Mixed methods. Concurrent verbal reports, text recall, prior knowledge, standardised reading test, tracking tool; qualitative analysis, descriptive statistical analyses.	919 propositions generated by 10 learners while reading hypermedia text: 829 were processing strategies and 90 were navigation strategies.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Ali (2013)	Critical thinking theory and the schema theory of reading comprehension	*N* = 10 Ages: 17 Language(s): EFL, Bahasa Melayu as L1 Language Level: N/A Academic Level: undergraduate Department: N/A Location: a university in Malaysia	To investigate differences in students' cognitive reasoning strategies in taking critical reading– thinking tests between Malay and English.	Mixed methods. Concurrent verbal reports, English Language Critical Reading Test, Bahasa Melayu Critical Reading Test; coding, inter-rater reliability, descriptive statistical analyses.	The students used inference as an over-arching strategy. Background knowledge and knowledge of formal schemata were important in the participants' reasoning.
Alsheikh (2011)	N/A	*N* = 3 Ages: N/A Language(s): Hausa (L1), French (L2), English (L3) Language Level: N/A Academic Level: graduate Department: N/A Location: Midwestern University in USA	To explore strategies used by multilingual readers when reading across three languages – Hausa, English, and French.	Qualitative. Background questionnaire, Survey of Reading Strategies (SORS), Set of expository reading passages in 3 languages, verbal report assessment for text comprehension, concurrent verbal reports; constant comparative analysis, descriptive statistical analyses.	Limited use of reading strategies in native language as compared with English and French; most proficient reader used greater variety of strategies.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Alsheikh and Mokhtari (2011)	N/A	$N = 90$ for surveys and 10 for verbal reports Ages: 17–47 Language(s): Arabic as L1, ESL Language Level: TOEFL score of 578 (mean) Academic Level: undergraduate and graduate Department: N/A Location: five Midwestern universities in the USA	To investigate advanced ESL learners' metacognitive awareness and reading strategies in Arabic as L1 and English as L2.	Mixed methods. A survey of reading strategies, reading task, concurrent verbal reports; descriptive statistical analyses, constant comparative analysis, interrater reliability.	The participants had similar preferences for reading strategies in Arabic and English, but reported using more support and problem solving strategies in English than in L1. In Arabic, the students used fewer strategies than in L2.
Kim (2011)	An intersection of Bakhtin's, Rosenblatt's, Gee's, and Friere's sociocultural perspective on learning	$N = 4$ Ages: 7–9 Language(s): ELLs Language Level: N/A Academic Level: 2nd and 3rd primary grade students Department: N/A Location: middle-class urban public elementary school, Southwestern USA	To examine how ELLs construct meaning.	Qualitative. Observations, interviews, concurrent and immediate retrospective verbal reports, reflective journals; qualitative analyses.	Themes emerging from data were related to ELLs' cultural perspective; ELLs' lived-through experiences; ELLs' efferent reading; ELLs' dialogic meaning construction; and ELLs' critical reading to learn.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Vorobel and Kim (2011)	Bakhtin's sociocultural perspective on interaction	$N = 7$ Ages: N/A Language(s): ELLs Language Level: low-intermediate to high-intermediate Academic Level: undergraduate Department: English Language Institute Location: urban research university	To explore the reading strategies used by college-level ESL learners for online L2 texts.	Qualitative. Concurrent and retrospective verbal reports, observation, semi-structured interviews; inductive and interpretive analyses.	Seven themes emerged from participants' online-reading strategy use: using hypermedia, using computer applications and accessories, dialoguing, setting up reading purposes and planning, previewing and determining what to read, connecting prior knowledge and experiences with texts and tasks.
Bengeleil and Paribakht (2004)	N/A	$N = 17$ Ages: 22–25 Language(s): EFL, Arabic Language Level: intermediate and advanced Academic Level: N/A Department: Medical Location: Libya	To investigate the effect of EFL learners' L2 reading proficiency on L2 lexical inferencing.	Mixed methods. Vocabulary Knowledge Scale (VKS) pretest, questionnaire, concurrent and retrospective verbal reports, VKS posttest administered after 2 weeks; descriptive statistical analyses.	Both groups used the same kinds of knowledge sources and contextual cues despite their reading proficiency level, with 1 exception. Advanced readers made more correct inferences than intermediate-level readers.

Table 1. (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Bowles (2011)	Schmidt's framework of attention and noticing hypothesis	*N* = 50 Ages: N/A Language(s): Naïve English speakers Language Level: N/A Academic Level: undergraduate Department: N/A Location: N/A	To examine effects of exposure to glosses on readers' noticing and acquisition of targeted vocabulary and text comprehension.	Quantitative. Pre/post test recognition tasks, pre/post test production tasks, comprehension task, concurrent verbal reports, tracking, debriefing questionnaire; coding, one-way ANOVA, repeated measures ANOVA.	Readers exposed to glossed text in both conditions reported noticing targeted words significantly more than readers exposed to same text with no glosses and experienced significant effects on comprehension of content.
Bowles and Leow (2010b)	Ericsson and Simon's framework for verbal reports	*N* = 45 Ages: N/A Language(s): ESL Language Level: Fifth-semester Spanish course Academic Level: N/A Department: N/A Location: N/A	To explore effects of type of verbalisation on L2 readers' comprehension, ability to produce old and new exemplars of targeted L2 structure, and time taken to complete the tasks.	Quantitative. Comprehension task, written production tasks, concurrent and concurrent introspective verbal reports; coding, one-way ANOVAs.	Nonmetalinguistic experimental group performed significantly better on comprehension than metalinguistic group. Both verbalisation groups spent significantly more time on task than silent control group.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Camps (1998)	Schmidt's framework of attention and noticing hypothesis	$N = 74$ Ages: N/A Language(s): native English speakers, Spanish as L2 Language Level: first-year Spanish speakers Academic Level: undergraduate Department: language laboratory Location: N/A	To determine whether L2 learners who notice target forms obtain better scores than those who do not; whether type of verbal report and time in course affects scores.	Mixed methods. Questionnaire, concurrent and retrospective verbal reports; descriptive statistical analyses, two-way ANOVAs, t tests, comparisons of t tests, coding, qualitative analyses.	Students who mentioned object pronouns and their agreement features in verbal reports did not obtain higher scores than those who did not mention same. For 2nd semester students, those who mentioned pronouns and features scored higher for both types of reports.
Chou (2012)	N/A	$N = 92$ Ages: N/A Language(s): EFL Language Level: B1 and B2 threshold level of the Common European Framework of Reference (CEFR) for Languages Academic Level: undergraduate Department: Chinese major, Journalistic English course Location: university in Taiwan	To investigate EFL students' strategy use when reading general and subject-specific texts in testing and non-testing contexts.	Mixed methods. Questionnaires, follow-up interviews, reading tasks, and retrospective verbal reports; Wilcoxon non-parametric paired sample signed-rank tests, descriptive statistical analyses.	EFL students' strategy use was similar when reading general and subject-specific texts. In both contexts, reading comprehension was determined by information-processing stages and accessibility of schemata. The participants adopted monitoring, regulating, and managing information when failing to use cognitive strategies.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Chun (2013)	N/A	$N = 23$ Ages: N/A Language(s): fluent in English, German as a Foreign Language, Language Level: second-year German course Academic Level: undergraduate Department: N/A Location: large university in Southern California	To investigate frequency with which learners consult internal glossary and external dictionary; to determine whether correlation exists between use of glossaries and learners' text comprehension.	Mixed methods. Reading tasks, summary task, tracking, concurrent introspective verbal reports, interviews; descriptive statistical analyses, correlation coefficients; t tests.	Learners looked up more words in internal than external dictionary and performed better on the measure of comprehension when there was access to both internal glossary and external dictionary. A significant positive correlation existed between total time on task and comprehension.
Comer (2012a)	N/A	$N = 12$ Ages: 20–30 Language(s): English as L1, Russian as L2 Language Level: intermediate, fifth-semester students of Russian Academic Level: N/A Department: N/A Location: a large Midwestern university in the USA	To investigate Russian as an L2 students' use of lexical inferencing and other strategies when reading informational texts.	Mixed methods. Background survey, reading task, concurrent verbal reports; coding, descriptive statistical analyses.	Russian as L2 learners most commonly used unexplained guesses, noting an unknown word, offering a sentence-level gist, and verbalising the intention to skip the unknown word. Morphological and phonological analyses were most frequently used inferencing strategies. The use of strategies varied depending on the type of informational text.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Comer (2012b)	N/A	$N = 12$ Ages: 21.5 (average age) Language(s): English as L1, Russian as L2, Language Level: intermediate, fifth-semester students of Russian Academic Level: N/A Department: N/A Location: a large Midwestern university in the USA	To investigate what grammatical knowledge learners of Russian use when reading informational texts.	Mixed methods. Reading task, concurrent introspective verbal reports, written recalls; coding, interrater agreement, descriptive statistical analyses.	The participants used few metalinguistic terms and varied in their comprehension of the text in Russian as well as in quality and quantity of grammatical comments.
Comer (2012b)	N/A	$N = 12$ Ages: N/A Language(s): English as L1, Russian as L2 Language Level: intermediate, fifth-semester students of Russian Academic Level: N/A Department: N/A Location: a large Midwestern university in the USA	To investigate students' use of Russian–English dictionaries while reading tests in Russian as L2.	Mixed methods. Reading task, concurrent introspective verbal reports; coding, descriptive statistical analyses.	Learners of Russian prioritised what words to look up in the bilingual dictionary based on their relevance. The use of dictionaries varied greatly on when reading informational texts in Russian.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Daalen-Kapteijns *et al.* (2010)	N/A	*N* = 16 Ages: 11–12 Language(s): N/A Language Level: selected on basis of test for Dutch vocabulary knowledge Academic Level: sixth grade Department: N/A Location: N/A	To identify activities focused on vocabulary knowledge that young students are capable of doing, given adequate circumstances and support.	Mixed methods. Reading tasks to figure out meanings of unknown words, concurrent verbal reports; qualitative scoring of verbal reports, *t* tests.	Higher verbal ability group gained significantly higher scores on 3 activities focused on vocabulary knowledge. Decontextualisation proved to be higher in students of higher verbal ability.
Dressler *et al.* (2012)	N/A	*N* = 12 Ages: N/A Language(s): ELLs, 8 Spanish–English bilinguals, 4 monolingual English-speakers Language Level: N/A Academic Level: fifth-grade students Department: N/A Location: Santa Cruz, California	To examine how Spanish-speaking ELLs use cognate knowledge to assign meaning to English words that are cognates, situations when this is most effective, and ways in which this is applied.	Qualitative. Interviews, concurrent verbal reports, reading tasks on 6 short passages with target cognates; coding, descriptive statistical analyses.	ELLs' use of the cognate was associated with strategy correct inferences for Spanish–English cognates. Spanish-speaking students were more likely to use cognate strategy as it had been taught. Cognate knowledge was the strategy most associated with accurate inferencing.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Gascoigne (2005)	Goodman's top-down and Brown's bottom-up models of reading	*N* = 16 Ages: average age of 20 Language(s): native English, French as a Second Language speakers Language Level: N/A Academic Level: N/A Department: N/A Location: university of Nebraska, Omaha	To provide insight into the role of various text-driven and reader-driven processes necessary for revisiting mental models of the L2 reading process.	Mixed methods. Reading tasks, concurrent verbal reports, recall task; scoring of idea units, *t* tests.	Learners collectively recalled 116 idea units with 15% representing main idea units, 11% representing high-level topics, and 13% representing mid-level ideas. 60% represented minor detail.
Geladari *et al.* (2007)	N/A	*N* = 32 Ages: N/A Language(s): bilingual Language Level: N/A Academic Level: 5th and 6th grade students Department: N/A Location: N/A	To examine difficulties, cognitive and metacognitive strategies encountered by bilingual students and the impact of language competence and type of bilingualism on comprehension.	Mixed methods. Expository and narrative reading tasks, concurrent and immediate retrospective verbal reports, semi-structured interviews; qualitative analyses, descriptive statistical analyses, chi-square and one-way ANOVA.	22 categories resulted from analysis of verbal report data. These categories comprised three thematic categories: (i) reading difficulties; (ii) cognitive strategies employment; and (iii) metacognitive strategies employment.

Table 1. (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Ghavamnia *et al.* (2010)	Oxford's taxonomy of language learning strategy	*N* = 8 Ages: 21–23 Language(s): EFL, Persian as L1 Language Level: proficient and less-proficient Academic Level: senior undergraduate students Department: TEFL major Location: a university in Iran	To investigate the type and frequency of strategies proficient and less-proficient Iranian EFL readers employ when reading expository texts.	Mixed methods. Reading task with multiple-choice comprehension test adopted from TOEFL test, concurrent verbal reports, semi-structured interviews; descriptive statistical analyses, intercoder reliability, coding.	Iranian EFL readers used 294 cognitive strategies with translating, highlighting, and summarising being employed most frequently, 86 metacognitive strategies (self-monitoring as most frequent). The participants also used compensation and memory strategies, but less frequently while social and affective strategies were not used at all. There was a difference in the use of strategies between proficient and less-proficient L2 readers.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Ghonsooly and Barghchi (2013)	N/A	$N = 90$ in the first phase and 4 in the verbal protocol phase Ages: N/A Language(s): EFL Language Level: pre-intermediate and upper-intermediate Academic Level: N/A Department: N/A Location: a private English language school in Mashhad, Iran	To investigate relationship between reading anxiety and reading proficiency as well as between reading anxiety and learners' use of reading strategies.	Mixed methods. Reading task, concurrent and retrospective verbal reports, Foreign Language Reading Anxiety Scale questionnaire; descriptive statistical analyses, Pearson Product-Moment Correlation.	No significant relationship was found between reading anxiety reading proficiency. Reading anxiety had an effect on EFL learners' reading style and their use of reading strategies.
Godfroid and Schmidtke (1967)	Schmidt's noticing hypothesis	$N = 29$ Ages: 19–28 Language(s): EFL, Dutch as L1 Language Level: advanced (B2 (upper intermediate) or C1 (lower advanced) levels of the Council of Europe's (2001) CEF) Academic Level: second-or third-year English language majors Department: TEFL major Location: Belgian university	To investigate the role of attention and awareness in receptive vocabulary learning.	Mixed methods. Reading task, eye-tracking, vocabulary posttest, post reading-task interview, retrospective verbal reports; coding, mixed- model regression analyses.	Attention to novel words was found to cause awareness, which itself turned out to be a strong predictor of vocabulary learning.

Table 1. (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Godfroid and Spino (2013)	N/A	$N = 102$ Ages: N/A Language(s): English as a Foreign Language Language Level: advanced, learning English for an average of 8 years Academic Level: 2nd and 3rd year undergraduate students Department: English major Location: Vrije Universiteit Brussel, Belgium	To examine the potential reactive effects of eye-tracking and verbal reports on ESL students' reading comprehension and vocabulary recognition.	Quantitative. Between-subject design, reading task, content and vocabulary recognition posttests, concurrent verbal reports, eye-tracking, language background questionnaire; superiority tests (generalised estimating equations logistic regression, Kruskal-Wallis test), equivalence tests, effect size.	Thinking-aloud and eye-tracking did not affect reading comprehension. Small positive effect was found in vocabulary recognition due to thinking-aloud.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Goo (2015)	Towse and Hitch's task-switching account of working memory versus working memory's resource-sharing nature	$N = 42$ Ages: N/A Language(s): English speaking learners Language Level: learners of Spanish as a foreign language Academic Level: N/A Department: N/A Location: American university	To explore the relationship between working memory and learner performance on comprehension and development of the Spanish immediate future.	Quantitative. Listening span task, operation span task, reading task, comprehension test, written production test, concurrent verbal reports; regression analysis, ANCOVA.	No direct evidence was found for the role of working memory capacity (WMC) in reading comprehension. The regression analysis showed a statistically significant result, which indicated that WMC predicted learner performance on the posttest (written production). Verbal reports did not have a negative effect on the development of learning the Spanish immediate future.
Gurses and Adiguzel (2009)	Cognitive Academic Language Learning Approach	$N = 18$ Ages: 18–23 years Language(s): French as a FL Language Level: N/A Academic Level: N/A Department: N/A Location: French Preparatory Program at Eskisehir Osmangazi University	To investigate French as L2 learners' use of reading strategies and the effects of reading strategies instruction based on the Cognitive Academic Language Learning Approach on LLs' comprehension.	Mixed methods. Reading comprehension achievement test, reading strategy scale, concurrent verbal reports; descriptive and inferential statistical analyses, Wilcoxon Signed Rank Test, coding	The reading strategies instruction based on the Cognitive Academic Language Learning Approach had a significant effect on LLs' reading comprehension. Because of reading strategies instruction, the participants used and varied the use of reading strategies more frequently.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Güvendir (2013)	N/A	$N = 30$ Ages: 19–21 years Language(s): Japanese ESL learners Language Level: N/A Academic Level: freshmen Department: English Language Teaching Location: Trakya University in Turkey	To investigate the use of verbal reports to identify factors which lead to anxiety when reading in FL.	Mixed methods. Reading task, concurrent verbal reports, interviews; content analysis, coding, intercoder reliability, Cohen's Kappa value, descriptive statistical analyses,	Twelve factors contributed to FL reading anxiety; metaphorical title, unknown vocabulary, exam, reading aloud, text length, time, teacher's questions, topic familiarity, linguistic structures, students' questions, coherence among paragraphs, and text type.
Hamada (2014)	N/A	$N = 5$ Ages: Average age of 21 Language(s): Japanese ESL learners Language Level: N/A Academic Level: N/A Department: N/A Location: mid-size university in the USA	To examine how L2 word-meaning inference strategies, variety of strategy use, and success with L2 word meaning inference change over time.	Qualitative. Inference sessions, comprehension check (summary), concurrent verbal reports; qualitative analyses.	Mean performance rate from Ben the highest, followed by Abby, Cathy, Ed, and Debby. Mean number of strategies used per word was highest from Abby, followed by Cathy, Debby, Ed, and Ben. Only Debby demonstrated considerable change. The highest number of strategies corresponded with the highest success rate in Debby and the lowest success rate in Abby and Cathy.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Hamada and Park (2009)	N/A	N = 3 Ages: 20–37 Language(s): ESL, Korean as L1 Language Level: intermediate–advanced Academic Level: undergraduate and graduate Department: architecture, education, and mathematics. Location: a mid-size university in the USA	To investigate ESL learners' inferencing of new vocabulary with the focus on inference accuracy and strategy use.	Qualitative. Reading task, concurrent verbal reports; coding, interrater reliability.	ESL learners who showed higher inference accuracy had preference for global strategies, rather than the local, and consistently applied the same types of strategies. ESL learners with lower inference accuracy more frequently applied a wider variety of strategies.
He (2009)	Ames and Archer's goal theory	N = 57 Ages: N/A Language(s): N/A Language Level: similar levels of English proficiency Academic Level: college Department: N/A Location: Taiwan	To explore the relationship between goal types and adult EFL readers' strategy use and comprehension.	Mixed methods. Goal scale, reading proficiency test, concurrent and immediate retrospective verbal reports, retellings, reading comprehension test; qualitative analyses, MANOVA, stepwise multiple regression analyses, one-way ANOVA.	The performance goal was a negative predictor for the frequency of use of comprehension with individual paragraphs (CIP), comprehension across paragraphs (CAP) and monitoring/evaluating comprehension (MEC) strategies. The strong mastery, strong performance goal profile group used the CIS, CIP, CAP, and MEC strategies most often.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Hijikata *et al.* (2008)	N/A	*N* = 2 Ages: early 20s Language(s): Japanese as L1, EFL Language Level: TOEIC score of Participant 1 was 300, while the TOEFL–PBT score of Participant 2 was 580 Academic Level: undergraduate Department: School of Management Location: A private university in Japan	To investigate the interactions among L2 proficiency, reading strategies, and the rhetorical features of the papers when reading academic papers in EFL.	Mixed methods. Video observation, verbal reports, document analysis of participants' notes taken while reading, pre- and post-reading interview; descriptive statistical analyses.	The participants could not read academic papers efficiently owing to difficulties with recognising academic discourse. They focused on understanding figures and formulas in the text. The purpose of frequent use of local and global strategies varied and depended on L2 competence, background knowledge, and the participants' familiarity with the discipline-specific academic discourse.
Horiba (2013)	N/A	Experiment 1 *N* = 84 Ages: 18–21 Language(s): English as FL Language Level: TOEFL 430–450 Academic Level: 1st-year English course, undergraduate Department: English major Location: a university in Japan	To investigate if EFL learners' L2 text comprehension differ when reading for expression, image, and critique and to examine if the relations between L2 text comprehension and L2 proficiency and general comprehension skill differ when reading for expression, image, and critique.	Quantitative. Reading task, recall tasks; a two-way analysis of variance (ANOVA), descriptive statistical analyses, ANCOVA.	The EFL learners' amount of recall when reading for expression, image, and critique was similar while the amount of resource allocation to word analysis, reaction and evaluation, and self-monitoring turned out to differ between task conditions.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
		Experiment 2 *N* = 26 Ages: 19–25 Language(s): English as FL, Japanese as L1 Language Level: N/A Academic Level: Applied Linguistics course Department: N/A Location: a university in Japan	To investigate reading comprehension when EFL learners have different reading goals and to examine how the process of reading comprehension relate to the product of comprehension.	Quantitative. Reading task, concurrent verbal reports, recall task; descriptive statistical analyses, MANOVA.	The process of reading comprehension in L2 differed when EFL learners had different reading goals. There was a significant effect of the interaction between task condition and process level. No significant correlation was found between recall and the proportion of verbal reports on process level and category (except association).
Hu and Nassaji (2013)	N/A	*N* = 11 Ages: N/A Language(s): ESL Language Level: advanced with TOEFL score of at least 575 Academic Level: 2nd and 3rd primary grade students Department: N/A Location: university in Canada	To investigate the relationships between the ease of L2 learners' inferencing of word meanings from context, inferential strategies, and subsequent retention of the target words.	Mixed methods. Reading task, vocabulary pre-tests and post-tests, concurrent and introspective verbal reports; quantitative and qualitative analyses, descriptive statistical analyses, point-biserial correlation, coding, inter-rater checks.	There was a negative relationship between the ease of inferencing and retention of word meanings. The relationship between the type and frequency of use of inferential strategies and retention was found to be significant.

Table 1. (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Hu and Nassaji (2012)	N/A	*N* = 11 Ages: N/A Language(s): ESL Language Level: N/A Department: Economics and business majors Location: University in Canada	To investigate ESL students' inferential strategies and characteristics of successful and less successful inferencers.	Mixed methods. Reading task, concurrent verbal reports, multiple-choice reading comprehension test; quantitative and qualitative analyses, descriptive statistical analyses, coding, inter-rater checks.	Successful inferencers frequently adopted evaluation and monitoring strategies, textual and background knowledge, self-awareness, and repeated attempts when inferring the target word meanings.
Jahandar *et al.* (2012)	N/A	*N* = 32 Ages: 18–23 Language(s): EFL Language Level: upper intermediate Department: Economics and business majors Location: Parsian English Institute in Tonekabon, Iran	To investigate the effect of think-aloud method on EFL learners' reading comprehension proficiency.	Quantitative. Concurrent verbal reports, verbal recalls, interviews, pretest and posttest; Samples Independent *t* test, descriptive statistical analyses.,	Think-aloud method had significantly positive effect on improving EFL learners' reading comprehension.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Karimi and Alibakhshi (2014)	N/A	$N = 22$ Ages: 17–29 Language(s): EFL with Farsi as L1 Language Level: advanced Department: N/A Location: private language centre	To investigate L2 learners' reading strategies when reading for single-text comprehension vs. multiple-texts integration.	Quantitative. Test of the participants' prior knowledge of the topic of the texts, reading proficiency test, reading tasks, retrospective written verbal reports; Cronbach's Alpha reliability, coding, descriptive statistical analyses, independent samples *t* tests with Holm–Bonferroni correction, inter-coder reliability.	EFL participants used significantly more metacognitive reading strategies when reading for multiple-texts-integration task, compared to single-text-comprehension task. There was no significant difference between the two tasks in the 'language-oriented' category of strategies.
Karnal and Pereira (2015)	Anderson's classification of reading strategies	$N = 10$ Ages: N/A Language(s): EFL; Chinese, Japanese, Arabic, Portuguese, Spanish as L1s. Language Level: intermediate Department: English Language Institute, PhD graduates, undergrads, and 1 specialist Location: University of Pittsburgh	To investigate cognitive strategies during reading academic texts in L2 with and without Google translator.	Qualitative. Questionnaire on reading habits, reading task, concurrent verbal reports; coding.	There was no difference in strategies when EFL students read with and without Google translator. However, with the use of Google translator, the use of strategies was more frequent that led to more reading comprehension.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Kasemsap and Lee (2015)	Typologies of reading strategies by Sheorey and Mokhtari and Phakiti	$N = 121$ at the 1st stage and 24 at the verbal reports stage Ages: N/A Language(s): EFL, Thai as L1. Language Level: lower and higher-level English language proficiency Department: majoring in accountancy Location: a vocational college in Thailand	To investigate Thai EFL students use of reading strategies.	Quantitative. Questionnaire surveys, reading task, retrospective verbal reports, semi-structured interviews; coding, statistical analyses.	Higher-level students used more retrieval strategies and all subscales of strategies (except for memory strategies) than the students with lower EFL proficiency. Apart from that, the difference in the use of reading strategies between the higher and lower level English proficient students was not significant.
Ketabi *et al.* (2012)	Anderson's inventory of reading strategies	$N = 8$ Ages: 20–28 Language(s): EFL Language Level: N/A Department: Linguistics Location: an Iranian University	To investigate Persian EFL learners' use of cognitive strategies when reading a hypermedia text.	Mixed methods. Reading a hypermedia text, questionnaire, concurrent verbal reports, researcher's notes, semi-structured interviews; coding, descriptive statistical analyses.	The proficient group of EFL readers used mainly authoring and prior knowledge while less-proficient learners mostly employed paraphrasing, translation into L1, and referring to the dictionary.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Kim and Cha (2011)	The Cognitive Academic Language Learning Approach (CALLA); Phakiti's typology of reading strategies	$N = 4$ Ages: N/A Language(s): EFL Language Level: intermediate Academic Level: 3rd year Department: majoring in pre-pharm medical science Location: a South Korean university	To investigate changes which Korean EFL students' made when regulating their cognition during reading.	Quantitative. Interviews, reading tasks, concurrent verbal reports; coding, descriptive statistical analyses.	Korean EFL students' regulation processes over cognition during reading were different in terms of frequency. Specifically, the participants' regulation became more flexible and organised by the end of the strategy training.
Ko (2018)	N/A	$N = 106$ Ages: 19–21 Language(s): English class Language Level: intermediate or high intermediate level Academic Level: 2nd semester of freshman, undergraduate Department: reading course that met twice a week for 50 minutes Location: N/A	To determine whether reading comprehension and reading strategies are affected by gloss type and identify the type of gloss preferred by learners.	Mixed methods. Reading task, multiple choice reading test, questionnaire; concurrent verbal reports, qualitative analyses, descriptive statistical analyses, one-way ANOVA.	There was a significant difference between the L2 gloss condition and the no gloss condition. When the no gloss conditions were compared, those who read the text under the no gloss condition used far more strategies than counterparts. Readers preferred glossed material.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Kremmel *et al.* (2005)	N/A	2 studies reported: *N* = 418 Ages: 16.9 (mean age) Language(s): German-L1 background (86%) Language Level: B2 of CEFR Academic Level: penultimate year of secondary education Department: N/A Location: six provinces of Austria	To examine if phraseological knowledge contributes to the prediction of reading performance in EFL.	Quantitative. Reading tasks with reading comprehension test (multiple choice, note form, and two types of multiple matching), multiple choice syntactic knowledge test, vocabulary knowledge test with selected- and constructed-response formats, multiple choice phraseological knowledge test; descriptive statistical analyses and reliability values, componential model of FL reading.	Measuring phraseological knowledge raised the explained variance in the reading test scores from 69% to 75%. Phraseological knowledge is the strongest contributor to the prediction of reading performance in EFL, followed by vocabulary knowledge.
		N = 15 Ages: 17 Language(s): German Language Level: B2 of CEFR Academic Level: penultimate year of secondary education Department: N/A Location: Austria	To explore advance EFL readers' use of multi-word expressions in reading.	Mixed methods. Reading task, concurrent verbal reports, multiple choice phraseological knowledge test, reading comprehension test (true \false\justification format); descriptive statistical analyses.	Knowledge of multi-word expressions in isolation most likely results in the correct reading test score, targeting multi-word expressions in context. EFL students paraphrased, explicitly elaborated on, implicitly used, or ignored phrasal expressions in context. More attention to phraseological target units was associated with more correct answers.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Lee-Thompson (2015)	Bernhardt's constructivist model of reading	*N* = 8 Ages: N/A Language(s): native English speakers, Chinese language students Language Level: intermediate level proficiency Academic Level: N/A Department: N/A Location: N/A	To examine the reading strategies used by American readers of Chinese and the difficulties encountered in reading narrative and argumentative text.	Mixed methods. Reading task, concurrent and concurrent introspective verbal reports; qualitative analyses, descriptive statistical analyses.	12 bottom-up and 15 top-down strategies used in text comprehension. Common difficulties experienced by learners were vocabulary, orthography, grammar, and background knowledge.
Leow (2011)	Schmidt's framework of attention and noticing hypothesis	*N* = 74 Ages: N/A Language(s): N/A Language Level: 1ˢᵗ year in university Spanish language program. Academic Level: college, undergraduate Department: N/A Location: N/A	To examine the relationship between exposure to enhanced input and (a) reporting of noticing of targeted forms; (b) L2 readers' immediate intake; (c) immediate written production; and (d) comprehension of content information.	Quantitative. Comprehension task, written production task, multiple-choice recognition task, concurrent verbal reports; parametric *t* test, Pearson product-moment correlation.	Amounts of reported noticing were statistically similar for both groups. Significant correlations between reported noticing and recognition for both the enhanced and unenhanced group. No significant difference in comprehension scores between the two groups.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Leow *et al.* (2001)	Van Patten's model of input processing (IP), lexical preference principle (LPP)	*N* = 72 Ages: N/A Language(s): N/A Language Level: average of 60 hours formal exposure to Spanish, 5th-semester Spanish course Academic Level: undergraduate Department: N/A Location: N/A	To determine the effect of type of attentional condition on adult L2 reading comprehension.	Mixed methods. Reading task, comprehension assessment, multiple choice assessment, concurrent verbal reports; one-way ANOVA, qualitative analyses.	Type of attentional condition had a differential effect on reading comprehension. There was no direct correlation between average comprehension scores and percentage of participants processing targeted items deeply.
Leow and Morgan-Short (2008)	N/A	*N* = 77 Ages: N/A Language(s): N/A Language Level: 1st-year Spanish program Academic Level: college level undergraduate Department: N/A Location: N/A	To examine the effect of thinking aloud on adult readers' comprehension, intake, and controlled written production.	Quantitative. Reading task, comprehension task, multiple-choice recognition task and fill-in-the-blank task, concurrent verbal reports; parametric *t*-tests, non-parametric *t* tests.	Thinking aloud while performing a reading task did not have a detrimental effect on adult readers' comprehension. Thinking aloud did not have a detrimental effect on adult readers' intake and controlled written production.

Table 1.　(Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Li and Suen (2012)	Differential skill functioning approach	N = 10 ESL learners for verbal reports; 522 ESL learners with an East Asian language background and 147 with a Romance language background Ages: N/A Language(s): ESL, Chinese, Spanish, Korean, Japanese, French, Romanian, Italian, and Portuguese as L1s Language Level: N/A Academic Level: N/A Department: N/A Location: N/A	To examine differences between a group with an East Asian language background and one with a Romance language background in terms of reading subskills.	Quantitative. Reading task, Michigan English Language Assessment Battery reading test, concurrent verbal reports; coding, Fusion Model of cognitive diagnostic modeling, logistic regression techniques, descriptive statistical analyses.	The two language groups differed significantly in terms of vocabulary and connecting and synthesising. Not a language groups but gender turned out a significant factor for syntax.
Li and Suen (2015)	Cross-linguistic transfer theory	N = 7 Ages: N/A Language(s): ESL, Chinese, Spanish, and Portuguese as L1 Language Level: N/A Academic Level: N/A Department: N/A Location: two adult ESL programs in a city on the East Coast of the USA	To investigate the difference between Chinese and Romance language–speaking ESL learners' word recognition during reading tests.	Qualitative. Reading section of the Michigan English Language Assessment Battery (MELAB) form E, concurrent and retrospective verbal reports, interviews; a constant comparison method.	Chinese ESL learners experienced more difficulties with word recognition during reading tests than Romance language-speaking L2 learners. They were found to heavily employ test-taking strategies to compensate for their gaps in vocabulary knowledge.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Li and Chun (2004)	N/A	*N* = 4 Ages: 18–20 Language(s): EFL, Chinese as L1 Language Level: completed 1st year English study Academic Level: undergraduate Department: Sociology and Computer Science Location: a Chinese university	To investigate differences between more proficient and less proficient Chinese EFL readers in terms of their use of metacognitive knowledge.	Qualitative. Reading task, concurrent and retrospective verbal reports, interviews, observations; coding, inter-rater reliability.	Vocabulary turned out to be a challenge for both proficient and less proficient readers. Less proficient readers preferred a bottom-up approach to reading while successful readers—a top-down approach.
Mahmoudi (2013)	N/A	*N* = 115 at the initial stage, 10 for verbal reports stage Ages: 19 20 Language(s): EFL, Iranian as L1 Language Level: more proficient and less proficient readers Academic Level: freshmen Department: majoring in electronics Location: Babol University of Technology, Mazandaran, Iran	To investigate how LLs with different proficiency levels monitor reading comprehension of academic texts.	Mixed methods. A standard reading test, concurrent verbal reports, error detection and retrospective questions; descriptive statistical analyses, *t* tests, coding, inter-rater coding.	More proficient EFL readers were found to use comprehension monitoring more often and more successfully than less proficient readers.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Mohammed *et al.* (2003)	VanPatten's models of input processing and Laufer and Hulstijn' involvement load hypothesis	*N* = 6 Ages: N/A Language(s): EFL Language Level: intermediate Academic Level: N/A Department: N/A Location: different universities in Iraq	To investigate a potential methodological issue called reactivity which might affect LLs' attention to form/meaning in various kinds of attentional conditions of task-induced involvement.	Qualitative. Reading task, concurrent verbal reports, experimental tasks; coding; descriptive statistical analyses.	Reading tasks with different involvement load caused different kinds of noticing.
Morgan-Short *et al.* (2009)	VanPatten's primacy of meaning principle	*N* = 308 Ages: 20.7 years (average) Language(s): Spanish as L2, English as L1 Language Level: 3rd semester Spanish program Academic Level: college level undergraduate Department: N/A Location: a Midwestern university	To investigate if simultaneous attention to form and meaning interferes with Spanish as L2 learners' reading, if there is a relationship between depth of processing and written L2 comprehension, and to examine reactivity of verbal reports.	Quantitative. Reading task, reading comprehension test, concurrent verbal reports; chi-square analysis, coding, a two-way ANOVA, Pearson's two-tailed correlation, descriptive statistical analyses.	Spanish as L2 learners who processed grammatical or lexical forms when reading achieved better comprehension. Verbal reports were found to have minimal effect and did not compromise the validity of the study.

Table 1. (Continued)

Study	Theoretical perspective	Participants/language/context	Research questions/purpose statement	Research methods. Data collection and analysis	Key findings
Nassaji (2015)	Pressley and Afflerbach's model of good strategy use	$N = 21$ Ages: N/A Language(s): Arabic, Chinese, Persian, Portuguese, Spanish as L1s; ESL Language Level: 12 week intermediate Academic Level: N/A Department: ESL program Location: Canada	To determine how successfully intermediate ESL learners infer word meanings from context in a reading text; the strategies and knowledge sources they used to do so.	Qualitative. Reading task, concurrent and immediate retrospective verbal reports; qualitative analyses.	Students used general knowledge of the world most frequently and depended heavily on this knowledge when inferring word meanings from context. The strategies learners used included repeating, verifying, monitoring, self-inquiry, and analysing.
Nassaji and Hu (2012)	Involvement Load Hypothesis	$N = 32$ Ages: N/A Language(s): Chinese Language Level: paper-based TOEFL score of at least 575 Academic Level: N/A Department: economics and business major Location: Canada	To investigate the relationship between task-induced involvement load and ESL learners' use and effectiveness of lexical inferencing strategies and learning word meanings from context.	Quantitative. Reading task, concurrent verbal reports, Nation's Vocabulary Levels Test, Vocabulary Knowledge Scale; inter-rater reliability of the coding, a two-way repeated measure ANOVA, descriptive statistical analyses, a one-way ANOVA.	There is a significant relationship between the degree of task-induced involvement load, learners' use of lexical inferencing strategies, and subsequent retention of successfully inferred words.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
O'Donnell (2011)	N/A	*N* = 197 Ages: N/A Language(s): N/A Language Level: late beginners or early intermediate LLs, 4th semester Spanish course Academic Level: undergraduate Department: N/A Location: N/A	To compare comprehension scores, ability to recognise words, and identify lexical items glossed in Spanish among readers of L3 literary text.	Quantitative. Reading task, comprehension recall assessment, concurrent verbal reports, assessment of vocabulary recognition; descriptive statistical analyses, *t* tests.	The amount of information that readers of elaborated text version recalled proved significantly greater than that of unmodified versions. Readers of elaborated versions performed better than the readers of unmodified versions.
O'Donnell (1996)	N/A	*N* = 18 Ages: N/A Language(s): N/A Language Level: beginning–intermediate Academic Level: undergraduate students Department: N/A Location: a large research university in the Midwestern USA	To examine intermediate L2 Spanish readers' use of marginal glosses to aid reading comprehension in print medium.	Mixed methods. Reading task, concurrent verbal reports; coding, descriptive statistical analyses.	The participants used glosses 54% of the time for facilitation of their reading comprehension. The misapplication of glosses was due to the formatting and placement of glosses and learners' interaction with them.

Table 1.　(Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Paribakht (2006)	N/A	*N* = 20 Ages: N/A Language(s): Farsi Language Level: high intermediate Academic Level: undergraduate Department: English majors Location: several universities in Iran	To examine the relationship between L1 lexicalisation of the concepts represented by the L2 target words and learners' inferencing behaviour while reading English texts.	Mixed methods. Vocabulary knowledge scale, vocabulary levels test, concurrent introspective verbal reports; qualitative analyses, descriptive statistical analyses.	A variety of knowledge sources from different levels of the language system were identified. Participants attempted to infer a greater percentage of non-lexicalised than lexicalised target words.
Park *et al.* (2012)	Hacker's self-regulated reading and Spiro's theory of cognitive flexibility	*N* = 7 Ages: N/A Language(s): ESL, Korean and Mandarin as L1 Language Level: advanced Academic Level: graduate Department: business, science majors (biology and chemistry), and education Location: US research university in the Midwest	To investigate L2 readers' information-seeking strategies and decision-making processes when reading online.	Qualitative. Pre-reading questionnaire and interview, reading task, comprehension test, observation, concurrent verbal reports, post-reading interviews; thematic coding analysis.	L2 readers actively used prior knowledge of the structure of offline and online resources when reading. The participants followed a recursive pattern of self-regulated reading strategies.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Radzi and Aziz (2014)	Schema theory	$N = 1$ Ages: N/A Language(s): ESL Language Level: N/A Academic Level: N/A Department: N/A Location: Universiti Teknologi MARA (Perlis), Malaysia	To investigate the role of content schemata in L2 learner's reading comprehension of content-familiar and unfamiliar texts.	Qualitative. Reading task, concurrent verbal reports. Interviews; coding.	Content schemata was found to be crucial for L2 learner's reading comprehension.
Rossomondo (2002)	Van Patten's model of IP, LPP	$N = 140$ Ages: N/A Language(s): one semester accelerated Spanish course at a Language Level: N/A Academic Level: N/A Department: N/A Location: large American university	To examine the how lexical temporal indicators (LTIs) affect comprehension and input processing.	Quantitative. Reading task, screening test, questionnaire, comprehension test, form-production test, concurrent verbal reports; one-way ANOVA.	There was greater comprehension in both text interaction formats for participants who interacted with the LTI passage.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Rott (2013)	Craik & Lockhart's levels of processing depth theory	*N* = 10 Ages: N/A Language(s): native English speakers learning German as a foreign language Language Level: N/A Academic Level: N/A Department: N/A Location: large public university in the USA	To explore why certain vocabulary interventions are more facilitative for word learning than others.	Mixed methods. Reading task, text comprehension, strategy use, concurrent verbal reports; qualitative analyses, descriptive statistical analyses.	Participants used a small variety of strategies. Word processing strategies were categorised as metacognitive word processing behaviours and semantic elaboration.
Rott and Gavin (2008)	Conceptual replication of Goldman *et al.* (2012)	*N* = 19 Ages: N/A Language(s): native English speakers learning German as a foreign language Language Level: intermediate (2nd year) and advanced (3rd and 4th year) Academic Level: undergraduate 2nd–4th year language and culture courses Department: N/A Location: midwestern public university	To explore interaction between reading multiple texts in German as a FL and processes of sense making, monitoring, and evaluation when reading and researching online.	Mixed methods. Reading and researching task, concurrent verbal reports, writing-an-essay task, reliability assessment of the texts/sites; qualitative analyses, descriptive statistical analyses, multiple *t* tests, an independent samples *t* tests, effect size, Cohen's Kappa.	Significantly more reading strategies were used by better learners than the poorer ones. The better learners engaged in more self-explanation and monitoring, skipped more unfamiliar vocabulary.

Table 1. (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Seng (1990)	Vygotsky's sociocultural perspective on learning	$N = 46$ Ages: N/A Language(s): enrolled in English course, ESL students Language Level: N/A Academic Level: undergraduate Department: N/A Location: N/A	To explore the use of think-alouds in a collaborative environment.	Quantitative. Reading task, concurrent verbal reports; ANCOVA.	The results showed that students in the experimental group obtained higher reading comprehension scores than their counterparts in the comparison group after the instruction with think-alouds.
Shin and Riazantseva (2013)	N/A	$N = 3$ Ages: 29–38 Language(s): EFL, Korean as L1 Language Level: novice-advanced Academic Level: graduate Department: Political Science, Electrical Engineering, Law School Location: a large university in the Northeastern USA	To explore Korean EFL readers' construction of meaning during reading and mediation of this process by learners' identities.	Qualitative. Reading task, concurrent verbal reports, semi-structured interviews, and participant written responses; coding.	The findings show the participants' positioning as native speakers of Korean, ESL learners, males, Christians, and graduate students. During reading a text and constructing meaning in ESL, Korean participants' L1 cultural identity emerged while their L1 culture informed and enhanced their comprehension of the plot of the text.

Table 1. (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Stevenson *et al.* (2007)	Walczyk's compensatory encoding model: orientation of processing; type of processing; domain of processing	*N* = 253 Ages: 13–14 Language(s): N/A Language Level: 3.5 years of EFL classroom instruction, low-intermediate Academic Level: 10th Grade 8 classes Department: N/A Location: six urban schools in the Netherlands	To examine 2 hypotheses about processing of global text content in second language reading.	Mixed methods. Concurrent verbal reports; qualitative analyses, ANOVAs.	There were differences in the proportional distribution of strategies across Dutch and EFL. The balance of processing for each of the 3 dimensions explored varied according to reader characteristics. The readers used a higher proportion of Language Oriented strategies and Regulatory strategies in EFL than in Dutch.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Tabataba'ian & Zabihi (1971)	O'Malley and O'Malley's (1990) model of language learning strategies	$N = 4$ Ages: 19–27 Language(s): EFL Language Level: studying in an upper-intermediate level Academic Level: undergraduate and graduate Department: majoring in Mechanical Engineering, Industrial Management Power Engineering, and Electrical Engineering. Location: College of Ferdowsi University in Mashhad, a city in north-eastern Iran.	To investigate the differences between reading strategies used by EFL learners when reading English for Specific Purposes and General Purpose English texts.	Quantitative. Reading task, concurrent verbal reports; coding, descriptive statistical analyses.	Unlike socioaffective strategies, cognitive strategies were used quite often in reading English for Specific Purposes and General Purpose English texts. EFL learners used background knowledge more often when reading English for Specific Purposes texts. Drawing on background knowledge as well as confirming the knowledge turned out to be important for reading English for Specific Purposes texts.

Table 1. (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Tabata-Sandom (2013)	N/A	*N* = 6 writers of grade readers (teachers of Japanese or former teachers of Japanese) and 14 learners of Japanese as a foreign language Ages: 21 (average) Language(s): Japanese as a foreign language Language Level: N/A Academic Level: N/A Department: N/A Location: main office of the Japanese Extensive Reading Research Group in a New Zealand university	To explore the intentions of writers of Japanese grade readers and Japanese as a foreign LLs' response to graded readers.	Qualitative. Focus group interview with writers of grade readers; a verbal report, the text comparison procedure and unstructured exit interviews with learners of Japanese as a foreign language; qualitative analyses, descriptive statistical analyses.	The writers of Japanese grade readers had a genuine communicative intent. While most participants had positive perceptions of grade readers owing to lexical simplification, learners of advanced level of proficiency in Japanese perceived grade readers in a negative way.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Upton and Lee-Thompson (1998)	Pressley and Afflerbach's model of good strategy use	*N* = 20 Ages: N/A Language(s): 10 native speakers of Chinese/10 native speakers of Japanese Language Level: intermediate advanced ESL and post-ESL Academic Level: graduate and undergraduate Department: N/A Location: N/A	To determine what role the L1 plays in the reading strategies of L2 readers.	Qualitative. Reading task, concurrent and retrospective verbal reports; qualitative analyses.	Intermediate ESL students tended to think about and process the L2 reading task using their L1 more frequently than advanced ESL students. L1 was turned on and actively used by L2 readers. Reliance on the L1 declined as proficiency in the L2 increased.

Table 1. (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Wang (2000)	N/A	*N* = 24 Ages: N/A Language(s): English, Chinese as L2 Language Level: beginning Academic Level: N/A Department: N/A Location: Midwestern USA	To identify strategies when reading with and without a pop-up dictionary and to explore reading comprehension gains due to the use of a pop-up dictionary over a short period of time.	Mixed methods. Reading tasks with writing recall tasks, concurrent verbal reports; coding, descriptive statistical measures, correlation coefficients, non-parametric analysis of Friedman's ANOVA, Wilcoxon signed-rank tests.	Regarding of the use of a pop-up dictionary, the Chinese as Foreign Language students used bottom-up strategies with L1, monitoring comprehension, and focusing on semantic cues. Without dictionaries, the students applied recognising characters, dividing word boundaries, and retrieving or guessing word meaning. Reading comprehension gains owing to the use of a pop-up dictionary over a short period of time were not significant.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Weil (2007)	Bernhardt's constructivist model of reading	*N* = 19 Ages: N/A Language(s): English, Korean Language Level: N/A Academic Level: Korean undergraduate students, Korean students from Intensive English Language Institute Department: N/A Location: USA	To examine the relationship between breadth of vocabulary, background experiences in learning English, and student skill in the reading of an academic text.	Mixed methods. Reading task, concurrent verbal reports, retelling task; coding, correlation coefficients.	The mean vocabulary score for the undergraduate students was greater than that for intensive English students. There was a moderate relationship between vocabulary size and total hours of high school English instruction.
Wesche and Paribakht (1995)	N/A	*N* = 10 Ages: N/A Language(s): ESL, French Language Level: intermediate–level with French L1 backgrounds Academic Level: N/A Department: N/A Location: out-of-class research sessions in Canadian university	To explore university ESL learners' responses to 5 different types of text-based vocabulary exercises.	Qualitative. Reading task, learner reflections, learner interviews, introspective concurrent, immediate retrospective, and delayed retrospective verbal reports; qualitative analyses.	For most learners, the majority of the tasks succeeded at least partially in eliciting attention to the relevant features of the target words. Most learners reported finding tasks interesting. Learning meanings of new words was incremental and involved multiple exposures.

Table 1. (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Yang (2013)	N/A	$N = 20$ Ages: N/A Language(s): Mandarin Chinese as L1, EFL Language Level: intermediate Academic Level: N/A Department: College of Engineering and Management Location: Taiwan	To investigate the status of reading strategies and comprehension monitoring strategies in reading.	Qualitative. Reading task, concurrent and retrospective verbal reports; qualitative analyses.	Readers used reading strategies and comprehension monitoring strategies to aid their reading and interpretation. Readers with insufficient language knowledge adopted reading strategies to solve problems. EFL readers were equipped with knowledge of comprehension monitoring.
Yanguas (2010)	N/A	$N = 94$ Ages: N/A Language(s): Spanish Language Level: N/A Academic Level: last semester of foreign language requirement Department: N/A Location: small private university, Northeastern Seaboard	To investigate the effects that different types of multimedia glosses have on text comprehension and vocabulary learning of computerised reading passages.	Qualitative. Reading task, concurrent verbal reports, pre-test– post-test production tasks, pre-test–post-test recognition tasks, multiple-choice comprehension task; coding, ANOVAs.	Participants exposed to multimedia glosses reported noticing the target vocabulary more than those in the control group. No significant effect on the type of gloss on production and recognition tasks.

Table 1. (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Yilmaz (2001)	N/A	$N = 10$ Ages: N/A Language(s): Turkish as a FL, Farsi, Arabic, Turkmen, Tajik, and Swahili as L1s Language Level: B1 (Threshold, pre-intermediate), based on CEFR for languages Academic Level: N/A Department: N/A Location: Turkish Language Teaching and Application Center (TÖMER) at Gaziosmanpasa University, Turkey	To investigate the extent of Turkish LLs' reference to Turkish derivational affixes and to identify knowledge sources the participants used when guessing the meaning of new words when reading.	Mixed methods. Pre-test, reading task, concurrent verbal reports, researcher's notes; qualitative analyses, coding, descriptive statistical measures.	Turkish LLs did not succeed in referring to affixes when guessing meaning of words from the context. The four knowledge sources identified were grammatical knowledge, discourse/text knowledge, morphological knowledge, and world knowledge.

Table 1.   (Continued)

Study	Theoretical perspective	Participants/language/ context	Research questions/ purpose statement	Research methods. Data collection and analysis	Key findings
Yin (2011)	Fukkink's model on the process of deriving word meaning from context	$N = 55$ Ages: 16–24 Language(s): EFL, Chinese as L1 Language Level: tertiary final, tertiary middle, tertiary initial, and senior secondary Academic Level: N/A Department: English language and literature Location: N/A	To investigate Chinese EFL learners' inferencing during reading.	Mixed methods. Reading task, introspective concurrent verbal reports, stimulated recall, interviews; coding, descriptive statistical measures.	There were numerous types of sequence of Chinese EFL learners' lexical inferencing. Some participants utilised context to guess the meaning of an unknown word when they could not find any morphological cues, others analysed parts of an unknown word, made a hypothesis, and looked at the word in the context.
Zhang *et al.* (2017)	Goodman and Smith's reader–versus text-driven reading; Anderson's information processing model	$N = 18$ Ages: N/A Language(s): English Language Level: N/A Academic Level: 4th–6th grade Department: N/A Location: Singapore	To examine the reading strategies used by Singaporean primary school pupils from a cognitive perspective.	Mixed methods. Reading task, concurrent verbal reports, interviews; coding via Nvivo software, t test, ANOVA.	Grade level did not show a strong relationship with ESL reading and language proficiency. More mature students used comprehension strategies more frequently and flexibly. Low proficiency students relied heavily on decoding. Primary students were less resilient and systematic in their metacognitive endeavours and cognitive strategies than adult learners.

Throughout this process, we acknowledged that our inability to access and include studies written in languages other than English posed a limitation to the scope of knowledge that we were able to gather regarding the use of verbal reports. Moreover, we were aware that the decision to choose studies published within the years 2000–2015 prevented us from reviewing seminal works published prior to this time period. We therefore drew partly from seminal works on L2 reading and verbal reports such as that of Chamot and El Dinary (1999), Jiménez *et al.* (1995, 1996), Kern (1994), Lomicka (1998) and Pritchard (1990). For instance, we were cognisant of the study conducted by Chamot and El Dinary (1999) that showed that the use of think-alouds in elementary foreign language immersion programmes helped to illustrate strategies that more-effective and less-effective third- and fourth-grade learners used to engage in reading and writing tasks in the classroom using the target language. It was also useful to acknowledge the work of Jiménez *et al.* (1995) who elicited think-alouds from two bilingual Latina speakers and a monolingual English speaker to describe and understand how the cognitive and metacognitive knowledge of a proficient bilingual reader who was Latina was reflected in comparison with the reading processes and strategies of a limited proficient bilingual reader and a proficient monolingual reader. We noted the research conducted by Kern (1994) that used intermediate French students' think-alouds to explain the extent to and conditions under which they engaged in mental translation while reading French texts, the role that translation played in the comprehension and the influence of reading ability in L2 on mental translation during reading. Additionally, we considered the research conducted by Lomicka (1998) that relied on think-alouds of college students to describe how full glossing during computerised reading in French promoted a deeper level of text comprehension, as well as the study conducted by Pritchard (1990) that drew from the think-alouds of adolescent American and Palauan students to describe the influence of cultural background on the ways in which reading is processed and comprehended. While none of these studies constituted the corpus of literature that functioned as the basis for the current review, we nonetheless referred to findings and insights from some of these works in our ensuing discussion of the literature.

Analysis

We began by collecting empirical studies for review. We first used organisational templates to categorise elements of each study based on well-known literature review studies (Fitzgerald, 1995; Dixon *et al.*, 2012). For each study, the following were outlined (see Table 1): (a) theoretical background; (b) information about participants (e.g. number of participants, age), language (e.g. language proficiency), and context; (c) research questions and purpose statement, (d) research method, data collection and analysis (retrospective, concurrent, introspective and think-aloud verbal protocols; qualitative/quantitative; strategies employed) and (e) key findings. We organised the information from the studies into relevant categories within the organisational template. After reading and rereading the data within the template, we used each research question to identify patterns inductively. Thematic analysis is, by its very nature subjective, but by clearly describing our analytical process, we demonstrate

how we arrived at conclusions (Ryan & Bernard, 2003). In the following sections we describe our three-stage analytical process.

Stage one. We made specific observations of the material outlined within our organisational templates (Patton, 2002) in iterative phases, using a new template to address each research question. We used different colors to underline words or phrases that represented topics that recur, indigenous typologies or categories, similarities and differences, and theory-related material (Ryan & Bernard, 2003). For example, we noted similarities and differences in methodological choice within studies, such as the type of verbal report used and participant characteristics such as age/academic level.

Stage two. Our second phase of analysis involved finding key words in context, noting word co-occurrence and then coding those terms (Ryan & Bernard, 2003). Through this process, we obtained categories different from those pre-established in our templates. For example, after considering the research questions and findings for all studies, we arrived at novel categories such as strategy use, comprehension, vocabulary and technology. Similarly, from other pre-established categories, such as methodology and theoretical framework, we derived subcategories, which we then examined to derive broader themes (Merriam, 2009).

Stage three. The third phase of analysis allowed us to confirm the trustworthiness of the derived themes (Patton, 2002). We worked collaboratively to determine categories and themes, but we disagreed about some areas. In these cases, we returned to the original studies to clarify our conceptions and arrive at more representative findings. We also periodically discussed patterns emerging from the study and confirmed our agreements.

Our detailed analytical process notwithstanding, we acknowledge that this review is influenced by our worldviews and interpretations of verbal reports as a tool in studies conducted within the L2 field. By acknowledging our stance as researchers who view social processes as important to language and literacy, we were able to attend systematically to studies that adopted multiple perspectives and therefore minimise the potential for bias in analysis and synthesis. Had other researchers approached the same material, they might have arrived at different categories or synthesised the material differently. Nonetheless, we hope that by offering our methodological approach for scrutiny we enable readers to trace the process through which we arrived at our findings.

Findings

To answer our research questions, we first synthesise patterns from findings observed in the empirical studies reviewed. We also highlight concerns emerging from our analysis of literacy research in which verbal reports were deployed.

Educational patterns and concerns

To address our first research question, we consider verbal reports in relation to: (a) content addressed in the studies; (b) theoretical approaches utilised; and (c)

methodological choices. We first present the findings and patterns as they relate to each of the above, followed by interpretive commentary in which we highlight concerns.

Content of empirical studies.   In this review of research that uses verbal reports as a methodological tool for exploring LLs' reading processes, we identified the following areas of focus: (a) strategy use, (b) comprehension, (c) vocabulary, and (d) technology (see Table 2 for the distribution of empirical studies across these categories). The categories are not mutually exclusive, and studies were organised based on the area of focus that researchers intended to explore.

Strategy use—Researchers in 39 of the research studies used verbal reports to explore various reading strategies (see Table 2). Specific reading strategies included monitoring (e.g. Stevenson *et al.*, 2007; Mahmoudi, 2014), inferencing (e.g. Hamada, 2009; Comer, 2012a), problem solving (Alsheikh, 2011) and comprehension (within individual sentences, within individual paragraphs and across paragraphs; He, 2008). Most researchers explored strategy use through concurrent reporting methods, while only a few deployed concurrent introspective (Chun, 2001; Hu & Nassaji, 2012; Comer, 2014), concurrent and immediate retrospective (He, 2008; Ghonsooly & Barghchi, 2012; Ghavamnia *et al.*, 2013), retrospective (Karimi & Alibakhshi, 2014; Kasemsap & Lee, 2015) or concurrent and retrospective strategies (Vorobel & Kim, 2011; Upton & Lee-Thompson, 2001; Yang, 2006).

Comprehension—In using verbal reports to explore comprehension, researchers primarily consider glosses or definitions/notes in the text (Ko, 2005; O'Donnell, 2012), attentional conditions ('i.e. requesting L2 readers to process a written text for meaning while paying attention to specified forms in the input;' Leow *et al.*, 2008, p. 683), exposure to enhanced/unenhanced input[1] (Leow, 2001) and the relative impact on comprehension. Across 36 studies in which verbal reports were used to explore comprehension, researchers primarily depended on concurrent reports, though there were a few studies which used concurrent introspective protocols (Paribakht, 2005; Comer, 2012b), concurrent and immediate retrospective protocols (Nassaji, 2003; Kim, 2011; Li & Suen, 2013), and concurrent and concurrent introspective protocols (Bowles & Leow, 2005; Lee-Thompson, 2008).

Vocabulary use—Twenty-five research studies in which verbal reports were used to explore vocabulary focused predominantly on ESL and English as a Foreign Language (EFL) learners (Nassaji, 2003; Akpinar, 2013) and focused on topics ranging from inferencing (Bengeleil & Paribakht, 2004; Hamada, 2009) to vocabulary size and its implications for readers (Weil, 2008). In most studies, the researchers used concurrent verbal reports, though a few deployed concurrent methods in conjunction with retrospective reflection (Camps, 2003; Bengeleil & Paribakht, 2004; Li & Suen, 2015), concurrent introspection (Weil, 2008), immediate retrospection and delayed retrospection (Wesche & Paribakht, 2000). One researcher used concurrent introspective methods to explore Chinese EFL learners' inferencing during reading (Yin, 2013) and another one used retrospective verbal reports to investigate the role of

Table 2.   Distribution of articles across the four categories: Research foci of studies

Research foci of studies	Study	Number of studies
Strategy use	Akyel and Erçetin (2009); Ali (2013); Alsheikh (2011); Alsheikh and Mokhtari (2011); Chou (2013); Chun (2001); Comer (2012a, 2014); Ghavamnia *et al.* (2013); Ghonsooly and Barghchi (2012); Griva & Mastrothanasis (2010); Gurses and Adiguzel (2013); Hamada (2009); Hamada and Park (2011); He (2008); Hijikata *et al.* (2013); Hu and Nassaji (2012, 2014); Jahandar *et al.* (2012); Karimi and Alibakhshi (2014); Karnal and Pereira (2015); Kasemsap and Lee (2015); Ketabi *et al.* (2012); Kim and Cha (2015); Ko (2005); Lee-Thompson (2008); Li and Suen (2013); Mahmoudi (2014); Nassaji (2003); Nassaji and Hu (2012); Park *et al.* (2014); Vorobel and Kim (2011); Rott and Gavin (2015); Stevenson *et al.* (2007); Tabataba'ian and Zabihi (2011); Upton and Lee-Thompson (2001); Wang (2014); Yang (2006); Zhang *et al.* (2007)	39
Comprehension	Bowles (2004); Bowles and Leow (2005); Chou (2013); Chun (2001); Comer (2012b); Gascoigne (2002); Geladari *et al.* (2010); Godfroid and Spino (2015); Goo (2010); Gurses and Adiguzel (2013); Güvendir (2014); Hamada (2009); He (2008); Horiba (2013); Jahandar *et al.* (2012); Kim (2011); Kremmel *et al.* (2015); Ko (2005); Lee-Thompson (2008); Leow (2001); Leow *et al.* (2008); Leow and Morgan-Short (2004); Li and Suen (2013); Li and Chun (2015); Morgan-Short, Heil, Botero-Moriarty & Ebert (2012); Nassaji (2003); O'Donnell (2009, 2012); Paribakht (2005); Radzi and Aziz (2013); Rossomondo (2007); Seng (2007); Shin and Riazantseva (2015); Tabata-Sandom (2013); Wang (2014); Yanguas (2009)	36
Vocabulary	Akpinar (2013); Bengeleil and Paribakht (2004); Bowles (2004); Camps (2003); Comer (2012a); Daalen-Kapteijns *et al.* (2001); Dressler *et al.* (2011); Godfroid and Schmidtke (2013); Godfroid and Spino (2015); Hamada and Park (2011); Hu and Nassaji (2012, 2014); Kremmel *et al.* (2015); Li and Suen (2013, 2015); Mohammed *et al.* (2015); Nassaji and Hu (2012); O'Donnell (2009); Paribakht (2005); Rott (2005); Weil (2008); Wesche and Paribakht (2000); Yanguas (2009); Yilmaz (2015); Yin (2013)	25
Technology	Akyel and Erçetin (2009); Bowles (2004); Chun (2001); Ko (2005); Park *et al.* (2014); Rott and Gavin (2015); Wang (2014); Yanguas (2009)	8

attention and awareness in receptive vocabulary learning (Godfroid & Schmidtke, 2013).

Technology—Researchers in eight studies relied on verbal reports to explore technology use in LLs' reading (see Table 2). Of these, four focused on glosses in computerised tasks with respect to vocabulary and comprehension (Chun, 2001; Bowles, 2004; Ko, 2005; Yanguas, 2009), while others focused on hypermedia environments

in relation to L2 readers' comprehension (Akyel & Erçetin, 2009; Park *et al.*, 2014; Rott & Gavin, 2015) and internal and external glossaries as well as dictionaries (Chun, 2001; Wang, 2014). The independent use of concurrent methods was common across all these studies.

Significant patterns and interpretive commentary—Our findings about the use of verbal reports present two challenges. The first relates to the *content* obtained from these reports and the second to the *conception of reading* adopted in empirical studies.

The heavy reliance on concurrent methods in exploring various aspects of reading appears to emphasise *what* reading processes occurred as opposed to *how* they occurred. It is, therefore, not surprising that despite the focus on multiple areas of reading (i.e. strategy use, comprehension), researchers (e.g. Ko, 2005; O'Donnell, 2009; Goo, 2010) continued to be primarily concerned with the product of reading, that is, the results of the reading task [e.g. assessment of text comprehension and vocabulary acquisition in O'Donnell (2009)]. In fact, very few researchers (e.g. Geladari *et al.*, 2010; Kim, 2011) were concerned with the reading process (i.e. how learners interact with text and context/construct during the reading act). This is curious, given the potential of protocol analysis to reveal ongoing processes. The limitations of a focus on *what* as opposed to *how* processes occur have been previously documented in the literature (Fitzgerald, 1995). Such significant patterns and concerns highlight the need for researchers to attribute greater importance to how LLs' reading processes occur.

In terms of the *conceptions of reading*, terms such as 'multiliteracies' and 'new literacies' have been used to denote the multiplicity of social factors inherent in reading and challenging theories that portray reading as a discrete skill (Street, 1995; New London Group, 1996; Paris, 2005). Like multiliteracies, the notion of 'new literacies' suggests 'literacy' as multidimensional in nature and comprises 'the skills, strategies, and dispositions necessary to successfully use and adapt to the rapidly changing information and communication technologies and contexts' (Leu *et al.*, 2004, p. 1572). Within this framework, LLs must not only complete tasks based on print-based text, but they must also use the Internet as a central technology for literacy. As noted previously, in a few of the studies reviewed verbal reports captured LLs' competencies with online tools, demonstrating how participants used computers, electronic devices, or particular websites. However, more should be done to reflect the multiplicity of social and technological contexts and the distributed nature of communication and digital technologies in 21st-century reading.

Theoretical approaches in empirical studies.   In the research reviewed, we distinguished between studies that explicitly used a theoretical framework to examine the constructs operationalised using verbal reports and those that did not. Of the 76 studies, 37 (approximately 49%) specified a theoretical framework or underlying model for research into the associated construct (see Table 3). The use of theoretical frameworks and models varied widely across the studies; however, most studies were grounded in or related to one of three trends: the *cognitivist trend* (e.g. Goo, 2010; Kim & Cha, 2015); *models of reading and language* (e.g. Gascoigne, 2002; Yin, 2013); and *sociocultural theory* (e.g. Seng, 2007; Kim, 2011). In addition, a few researchers depended on

Table 3. Theoretical frameworks in original studies

Theoretical frameworks	Study	Number of studies
Top-down, bottom-up models of reading	Gascoigne (2002)	1
Bernhardt's constructivist model of reading	Lee-Thompson (2008); Weil (2008)	2
Pressley and Afflerbach's model of good strategy use	Nassaji (2003); Upton and Lee-Thompson (2001)	2
Ericsson and Simon's framework for the use of verbal reports	Bowles and Leow (2005)	1
Schmidt's framework of attention and noticing hypothesis	Bowles (2004); Camps (2003); Godfroid and Schmidtke (2013); Leow (2001)	4
Stanovich's short-circuit effect; Goodman and Smith's reader-driven reading versus text-driven reading; Anderson's information processing model of comprehension	Zhang *et al.* (2007)	1
Working memory and its role in cognitive performance: resource sharing versus executive attention; inhibitory-based executive control; cognitive control; proactive control	Goo (2010)	1
Compensatory encoding model: Orientation of Processing (language oriented; content oriented); type of processing (regulatory, cognitive, cognitive-iterative); domain of processing (below-clause level, clause level, above-clause level)	Stevenson *et al.* (2007)	1
Primacy of meaning principle; Van Patten's model of input processing; Van Patten's lexical preference principle (LPP)	Leow *et al.* (2008); Mohammed *et al.* (2015); Morgan-Short *et al.* (2012); Rossomondo (2007)	4
Craik & Lockhart's levels of processing depth theory	Rott (2005)	1
Goal theory	He (2008)	1
Sociocultural perspective on learning	Kim (2011); Vorobel and Kim (2011); Seng (2007)	3
Hacker's self-regulated reading and Spiro's theory of cognitive flexibility	Park *et al.* (2014)	1
Cross-linguistic transfer theory	Li and Suen (2015)	1
Oxford's taxonomy of language learning strategy	Ghavamnia *et al.* (2013)	1
Cognitive academic language learning approach	Gurses and Adiguzel (2013); Kim and Cha (2015)	2
Anderson's inventory of reading strategies	Karnal and Pereira (2015); Ketabi *et al.* (2012)	2
Involvement load hypothesis	Mohammed *et al.* (2015); Nassaji and Hu (2012)	2
O'Malley and Chamot's model of language learning strategies	Tabataba'ian and Zabihi (2011)	1
Differential skill functioning approach	Li and Suen (2013)	1

Table 3.   (Continued)

Theoretical frameworks	Study	Number of studies
Critical thinking theory and the schema theory of reading comprehension	Ali (2013)	1
Schema theory	Radzi and Aziz (2013)	1
Typologies of reading strategies by Sheorey and Mokhtari and Phakiti	Kasemsap and Lee (2015); Kim and Cha (2015)	2
Fukkink's model on the process of deriving word meaning from context	Yin (2013)	1

L1 models of reading and language as a basis for research into L2 reading processes (e.g. Gascoigne, 2002; Nassaji, 2003; Zhang *et al.*, 2007). We coded studies under the 'models of reading and language' theme if the author(s) employed a specific conceptual framework (e.g. Pressley & Afflerbach's (1995) model of good strategy use), rather than a broader theoretical lens. In doing so, we acknowledge a degree of overlap across these categories.

Cognitivist trend—Among studies in which verbal reports were deployed from a cognitivist perspective, Schmidt's framework of attention and noticing was used as a basis for research in four studies (e.g. Bowles, 2004; Godfroid & Schmidtke, 2013). Other cognitivist perspectives—such as the primacy of meaning principle (e.g. Leow *et al.*, 2008; Mohammed *et al.*, 2015), goal theory, and Ericsson and Simon's (1984) framework for the use of verbal reports (Bowles & Leow, 2005)—were used in other studies (e.g. He, 2008; Kim & Cha, 2015). In studies with a cognitivist orientation, most researchers (75%) tended to use concurrent methods (e.g. Rossomondo, 2007; Goo, 2010).

Sociocultural trend—Researchers in three studies (e.g. Seng, 2007; Kim, 2011; Vorobel & Kim, 2011) who adopted a sociocultural approach to research using verbal reports relied primarily on Bakhtinian and Vygotskian notions of the sociocultural nature of learning (Seng, 2007; Kim, 2011; Vorobel & Kim, 2011). These researchers relied on concurrent methods independently or on concurrent methods in conjunction with retrospective reports.

Trends with reading and language models—Researchers used both L1 and L2 reading and language models to support the use of verbal reports in reading studies. Among the L1 models were the top-down/bottom-up models of reading (Goodman, 1968; Brown, 1998; Gascoigne, 2002), Pressley and Afflerbach's (1995) model of good strategy use (Upton & Lee-Thompson, 2001; Nassaji, 2003), Stanovich's (2000) short-circuit effect, Goodman's (1967) and Smith's (1971) reader-driven versus text-driven reading, and Anderson's (2005) information-processing model of comprehension (Ketabi *et al.*, 2012; Karnal & Pereira, 2015). When these models were used, researchers deployed concurrent reports independently, as well as in conjunction with

immediate retrospective, retrospective, and introspective reports (e.g. Gascoigne, 2002; Lee-Thompson, 2008). In the cases where L1 reading and language models were used as a framework for understanding the reading process, concurrent verbal reports tended to be used independently as well as in conjunction with other forms of reporting (e.g. Gascoigne, 2002; Zhang *et al.*, 2007).

Among the L2 models drawn upon were the O'Malley and Chamot's model of language learning strategies (Tabataba'ian & Zabihi, 2011), Bernhardt's constructivist model of reading (Lee-Thompson, 2008; Weil, 2008), Schmidt's framework of attention and noticing hypothesis (Leow, 2001; Camps, 2003; Bowles, 2004; Godfroid & Schmidtke, 2013), Primacy of Meaning Principle (Rossomondo, 2007; Leow *et al.*, 2008; Morgan-Short *et al.*, 2012; Mohammed *et al.*, 2015), cross-linguistic transfer theory (Li & Suen, 2015), Oxford's taxonomy of language learning strategy (Ghavamnia *et al.*, 2013), and Cognitive Academic Language Learning approach (Gurses & Adiguzel, 2013; Kim & Cha, 2015). When these models were used, researchers relied primarily on concurrent reporting (e.g. Leow, 2001; Bowles, 2004; Rossomondo, 2007; Leow *et al.*, 2008; Tabataba'ian & Zabihi, 2011; Morgan-Short *et al.*, 2012; Mohammed *et al.*, 2015) and less so on concurrent and concurrent introspective (Lee-Thompson, 2008; Weil, 2008), concurrent and retrospective (Camps, 2003), and retrospective (Godfroid & Schmidtke, 2013) protocols.

Across the theoretical perspectives and approaches deployed, the format in which findings were presented varied considerably. This variation notwithstanding, we observed two patterns. First, in cognitively oriented studies, the tendency was to report quantitatively oriented results—that is, results that focused on statistical significance (e.g. Goo, 2010); numbers of students, strategies, concepts attained (Lee-Thompson, 2008); averages (Weil, 2008); and relationships (Leow, 2001; Zhang *et al.*, 2007). Conversely, when studies adopted either a sociocultural perspective or a reading and/or language model to examine the reading process (Lee-Thompson, 2008; Vorobel & Kim, 2011), the analyses tended to be qualitative—that is, researchers focused on patterns such as difficulties experienced by learners in reading and the types of strategies used in particular instances.

Interpretive commentary—Based on these insights, it appears that cognitivist orientations to studies continue to be prevalent. However, sociocultural and L2 models are also emerging as a basis for examining verbal reporting of reading processes. Beyond this, there is evidence of a preponderance of concurrent measures in most of the studies, whether independently or in conjunction with other verbal reporting mechanisms (i.e. introspective, retrospective). This phenomenon calls attention to the role of language proficiency, instructions in verbal reporting, and reactivity as relates to verbal reporting for L2 readers.

Cognitivist conceptions of verbal reports are derived from information-processing theory, in which verbalisation is presumed to act as a 'window into the minds of learners' (Bowles, 2010a, p. 2). With such perspectives, verbal reports are presumed to represent cognition; as a consequence, considerable attention is devoted to ensuring the use of measures that reduce the reactive effects (i.e. changes verbal reports might make in cognitive process) accompanying such reports. Despite the reliance on an information-processing model of verbal reports in the majority of studies for which a

reported theoretical model was provided (approximately 49%), researchers paid little to no attention to reactivity, that is, 'the positive or negative effects of think-alouds on learners' processing of the language while performing the task' (Leow *et al.*, 2008, p. 676). In fact, the few studies in which the reactive effects of verbal reports were attended to represent those geared specifically toward an understanding of reactivity as a construct and were not necessarily focused on the reading process or on the impact of this for linguistic variation in participant reading per se (e.g. Camps, 2003; Leow & Morgan-Short, 2004; Bowles & Leow, 2005; Goo, 2010). In her numerous investigations into reactivity and verbal reporting, Bowles (2008, 2010a) notes that variables such as L2 proficiency level and explicitness of instruction accounted for reactivity in given tasks (Bowles, 2010a). Yet we found very few instances in which researchers acknowledged the role of participants' L2 proficiency or researcher instructions in their findings. The significance of an inattention to such issues central to the cognitivist perspective becomes even more apparent when considering recent theoretical challenges to the cognitivist approach to verbal reporting measures.

While logical arguments have been made for the validity of verbal reports within a cognitive perspective, the evidence presented in the few studies where a sociocultural framework was used challenges the legitimacy of these arguments. In contrast to a cognitivist perspective, sociocultural approaches to verbal reports regard speech as a social construction and, therefore, question the view of verbal reports as a mere reflection of cognitive processes (Smagorinsky, 2011). Though no study within this review employed a sociocultural approach to the verbal reporting method per se, the few studies in which sociocultural approaches framed the reading process (i.e. Seng, 2007; Kim, 2011; Vorobel & Kim, 2011) involved a sociocultural interpretation of the protocols obtained from participant reports. That is, through the use of integrated concurrent, retrospective, and introspective verbal report methods, and even when concurrent measures were employed independently, researchers described the embeddedness of reading processes in social interactions. For instance, Seng (2007) observed how participants producing think-alouds as they read in a collaborative environment performed better on reading comprehension tests when compared with students who had not engaged in such tasks.

Nevertheless, whether researchers' deployment of verbal reports was based on a cognitivist or sociocultural approach, an even more significant challenge comes from the reading and language models used to frame studies in which the reporting occurred. As indicated in the findings, the use of monolingual reading theories was more prevalent than that of those that foregrounded L2 reading while relying on verbal reports. This is not surprising. In previous reviews of research on ESL learners, it has been suggested that they undergo 'substantively the same' cognitive reading processes observed in native speakers of English, allowing for latency with some facets of these processes for LLs (e.g. Fitzgerald, 1995, p. 180; Grabe, 2009). Despite this evidence, and while L2 reading continues to be heavily informed by L1 reading theory (Grabe, 2009; Kim, 2011), applying L1 reading and language models to L2 reading processes has been criticised owing to the lack of consideration to the cross-linguistic nature of L2 reading (Koda, 2007; Grabe, 2009; Kim, 2011). There is evidence that L1 reading and language models fail to consider the cross-linguistic features of L2 reading because they are based on English and tend to reflect English conceptions of

literacy (Grabe, 2009). The question therefore arises as to the expectations associated with reporting done by L2 learners when these learners' reading was analysed using assumptions applicable to monolingual learners. By failing to attend to cross-linguistic constraints evident in LLs' verbal reporting about text—the various ways in which language differences permeate the various reporting tasks in which readers engage— researchers could easily overlook components of processing specific to linguistic disparity.

The L2 reading and language models on which researchers relied for examining reading via verbal reporting appeared to address some of these concerns. For instance, Bernhardt's constructivist model of reading allowed her to examine the reading strategies used by American readers of Chinese and the difficulties encountered in reading narrative and argumentative text. In turn, the researcher identified 12 bottom-up and 15 top-down strategies used in text comprehension by Chinese language students as compared with native English speakers. Also focused on reading strategies, the Cognitive Academic Language Learning Approach allowed researchers to investigate French as L2 learners' use of reading strategies and the effects of reading strategies' instruction based on the LLs' comprehension. Findings from this study showed that as a result of reading strategies' instruction, the participants used and varied the use of reading strategies more frequently (Gurses & Adiguzel, 2013). In an instance where the focus was placed on word recognition, the use of cross-linguistic transfer theory showed that Chinese ESL learners experienced more difficulties with word recognition during reading tests than Romance language-speaking L2 learners and that they heavily employed test-taking strategies to compensate for their gaps in vocabulary knowledge (Li & Suen, 2015). Taken together, the emphasis on L2 reading and language models to examine the reading processes of learners in which verbal reports were used provide lenses unique to the linguistic affordances presented by these learners. Without these lenses, it is easy for researchers to miss the salient linguistic differences across languages as they become represented in reading. A reliance on both concurrent as well as on retrospective and introspective reporting (though limited) can provide clearer insights into the ways in which these learners' reading processes differ from that of their L2 peers during verbal reporting.

Methodological choices in empirical studies.   We identified the methodological concerns arising when verbal reports were deployed in studies of LLs' reading processes. We explore these concerns by examining studies with the following foci: (a) mixed methods (39 studies); (b) quantitative (20 studies), and (c) qualitative (18 studies) (see Table 4). In addition, we note the significant findings from studies dedicated to the exploration of verbal reports and how they functioned in understanding LLs' reading processes (see Table 5).

Mixed methods—The majority of researchers used mixed methods when deploying verbal reports to investigate LLs' reading processes. Within mixed-methods studies, researchers tended to use verbal reports to collect data concerning participants' reading processes, qualitatively code this data based on predefined models of strategy use (e.g. Akyel & Erçetin, 2009; Geladari *et al.*, 2010; Ketabi *et al.*, 2012) and/or other categories (e.g. Leow, 2001; Comer, 2012a) and subsequently use the categories to

Table 4.   Methodological constructs of original studies

Methodological construct	Study	Number of studies
Quantitative	Bowles (2004); Bowles and Leow (2005); Godfroid and Spino (2015); Goo (2010); Horiba (2013); Jahandar *et al.* (2012); Karimi and Alibakhshi (2014); Kasemsap and Lee (2015); Kim and Cha (2015); Kremmel *et al.* (2015); Leow (2001); Leow and Morgan-Short (2004); Morgan-Short *et al.* (2012); Li and Suen (2013); Nassaji and Hu (2012); O'Donnell (2009); Rossomondo (2007); Seng (2007); Tabataba'ian and Zabihi (2011); Yanguas (2009)	20
Qualitative	Alsheikh (2011); Dressler *et al.* (2009); Hamada (2009); Hamada and Park (2011); Karnal and Pereira (2015); Kim (2011); Li and Suen (2015); Li and Chun (2015); Mohammed *et al.* (2015); Nassaji (2003); Park *et al.* (2014); Vorobel and Kim (2011); Radzi and Aziz (2013); Shin and Riazantseva (2015); Tabata-Sandom (2013); Upton and Lee-Thompson (2001); Wesche and Paribakht (2000); Yang (2006)	18
Mixed methods (quantitative and qualitative)	Akpinar (2013); Akyel and Erçetin (2009); Alsheikh and Mokhtari (2011); Camps (2003); Ali (2013); Bengeleil and Paribakht (2004); Chou (2013); Chun (2001); Comer (2012a,b, 2014); Daalen-Kapteijns *et al.* (2001); Gascoigne (2002); Geladari *et al.* (2010); Ghavamnia *et al.* (2013); Ghonsooly and Barghchi (2012); Godfroid and Schmidtke (2013); Gurses and Adiguzel (2013); Güvendir (2014); He (2008); Hijikata *et al.* (2013); Hu and Nassaji (2012, 2014); Ketabi *et al.* (2012); Ko (2005); Kremmel *et al.* (2015); Lee-Thompson (2008); Leow *et al.* (2008); Mahmoudi (2014); O'Donnell (2012); Paribakht (2005); Rott (2005); Rott and Gavin (2015); Stevenson *et al.* (2007); Wang (2014); Weil (2008); Yilmaz (2015); Yin (2013); Zhang *et al.* (2007)	39

conduct further quantitative analyses (i.e. go from qualitative to quantitative; e.g. Paribakht, 2005; Akyel & Erçetin, 2009; Geladari *et al.*, 2010). For instance, the data reported in a number of these studies focused on areas such as *variety* (Paribakht, 2005; Rott, 2005), *averages* (Hamada, 2009), *proportional distribution* (Stevenson *et al.*, 2007), *quantities* (Lee-Thompson, 2008; Geladari *et al.*, 2010), *relationships* (Ghonsooly & Barghchi, 2012) and *comparisons* (Ko, 2005; Mahmoudi, 2014)—all of which were quantitatively oriented.

Sometimes researchers who used mixed methods deviated from this norm. In these less common cases, verbal reports served to generate categories for both quantitative and qualitative analyses in order to extend conceptual understandings of phenomena appearing in the findings. That is, they went from qualitative to quantitative and then reverted to qualitative. For example, Chun (2001); Gascoigne (2002) both

Table 5.   Verbal report methodologies of original studies

Verbal report methodology	Study	Number of studies
Concurrent verbal reports	Akpinar (2013); Akyel and Erçetin (2009); Ali (2013); Alsheikh (2011); Alsheikh and Mokhtari (2011); Bowles (2004); Comer (2012a); Daalen-Kapteijns *et al.* (2001); Dressler *et al.* (2011); Gascoigne (2002); Ghavamnia *et al.* (2013); Godfroid and Spino (2015); Goo (2010); Gurses and Adiguzel (2013); Güvendir (2014); Hamada (2009); Hamada and Park (2011); Hijikata *et al.* (2013); Horiba (2013); Hu and Nassaji (2014); Jahandar *et al.* (2012); Karnal and Pereira (2015); Ketabi *et al.* (2012); Kim and Cha (2015); Ko (2005); Kremmel *et al.* (2015); Leow (2001); Leow and Morgan-Short (2004); Leow *et al.* (2008); Li and Chun (2015); Mahmoudi (2014); Mohammed *et al.* (2015); Morgan-Short *et al.* (2012); Nassaji and Hu (2012); O'Donnell (2009, 2012); Park *et al.* (2014); Radzi and Aziz (2013); Rossomondo (2007); Rott (2005); Rott and Gavin (2015); Seng (2007); Shin and Riazantseva (2015); Stevenson *et al.* (2007); Tabataba'ian and Zabihi (2011); Tabata-Sandom (2013); Wang (2014); Yanguas (2009); Yilmaz (2015); Zhang *et al.* (2007)	50
Concurrent introspective verbal reports	Chun (2001); Comer (2012b, 2014); Paribakht (2005); Hu and Nassaji (2012); Yin (2013)	6
Concurrent and immediate retrospective verbal reports	Geladari *et al.* (2010); Ghonsooly and Barghchi (2012); He (2008); Kim (2011); Li and Suen (2013); Nassaji (2003).	6
Introspective concurrent, immediate retrospectiveand delayed retrospective verbal reports	Wesche and Paribakht (2000)	1
Concurrent and retrospective	Bengeleil and Paribakht (2004); Camps (2003); Li and Suen (2015); Vorobel and Kim (2011); Upton and Lee-Thompson (2001); Yang (2006)	6
Concurrent and concurrent introspective	Bowles and Leow (2005); Lee-Thompson (2008); Weil (2008)	3
Retrospective	Godfroid and Schmidtke (2013); Karimi and Alibakhshi (2014); Kasemsap and Lee (2015)	3

successfully employed concurrent verbal reports to code propositions from protocols, submitted these to statistical analyses and then qualitatively identified and derived salient points concerning participants under observation from these protocols. In both cases, the types of data derived from these studies were qualitative and provided insight into how and why learners engaged in processes.

Across mixed-methods studies, two patterns stood out. First, the number of participants varied significantly, ranging from two (Hijikata *et al.*, 2013) to 253 participants (Stevenson *et al.*, 2007). In two studies with more than 100 participants, independent concurrent reports were used (Ko, 2005; Stevenson *et al.*, 2007); in studies with fewer participants, researchers relied on integrated concurrent reports (e.g. Akpinar, 2013; Ali, 2013). This pattern is reasonable considering the data collection burden for large numbers of participants. Second, even when researchers analysed verbal reports qualitatively, they frequently used these results to generate descriptive statistical analyses (e.g. Lee-Thompson, 2008; Rott, 2005.

Quantitative studies—Twenty studies in this review involved solely quantitative analyses. In requiring participants to produce verbal reports, researchers in these studies tended to assign participants randomly to control and experimental conditions (e.g. Yanguas, 2009; Goo, 2010). Common analyses included ANOVAs (e.g. Yanguas, 2009; Horiba, 2013) and *t* tests (Leow, 2001; Karimi & Alibakhshi, 2014); less frequently, statistical analyses included correlations (Leow, 2001; Morgan-Short *et al.*, 2012) and ANCOVAs (Seng, 2007; Horiba, 2013). In experimental designs, comparisons were usually performed—for example, between various types of verbal reports and the observed effects of varied reading constructs (e.g. Seng, 2007; Goo, 2010) and types of glosses (e.g. Bowles, 2004; O'Donnell, 2009; Yanguas, 2009). Participants were generally expected to perform reading, production, comprehension, and recognition tasks in conjunction with concurrent verbal reports.

Significant variation occurred in the number of participants in quantitative studies. A few studies involved 30–50 participants (e.g. Goo, 2010; Nassaji & Hu, 2012), while several others included somewhat larger numbers, 70–80 participants (e.g. Leow, 2001; Horiba, 2013). In several studies, quantitative studies involved more than 100 participants (e.g. Rossomondo, 2007; Godfroid & Spino, 2015). In two studies, only four individuals participated (Tabataba'ian & Zabihi, 2011; Kim & Cha, 2015). Across quantitative studies, researchers relied mostly on independent concurrent reports irrespective of the quantitative method and the number of participants involved.

Qualitative orientation—In 18 studies, solely qualitative analyses were used on verbal report data. Researchers in these studies used verbal reports to obtain qualitative information, such as characteristics of readers' processes (Chun, 2001; Gascoigne, 2002) and readers' uses of different reading strategies in different languages (Hamada, 2009; Alsheikh, 2011; Vorobel & Kim, 2011).

Several of the studies involved coding information from verbal reports in conjunction with descriptive statistics, followed by narrative explanations of patterns emerging from the findings (e.g. Alsheikh, 2011; Dressler *et al.*, 2011). Others relied strictly on inductive qualitative analyses (e.g. Yang, 2006; Kim, 2011; Vorobel & Kim,

2011). For example, Upton and Lee-Thompson (2001) coded the data from concurrent and retrospective protocols based on Pressley and Afflerbach's (1995) three reading strategy types (i.e. identifying, monitoring and evaluating) and then created classifications of these types based on patterns found in the reports. Nassaji (2003) approached his study quite differently, assigning values of 0, 1, or 2 to students' verbal responses to represent their success in inferring word meanings from context. Subsequently, he used the ratings to describe patterns in students' verbal reports and the circumstances under which certain strategies often became apparent.

The number of participants involved in qualitative studies ranged from 1 to 21. Very few studies involved fewer than three participants (e.g. Alsheikh, 2011; Kim, 2011; Radzi & Aziz, 2013) and most studies involved participants toward the higher end of this range (e.g. Nassaji, 2003; Yang, 2006; Vorobel & Kim, 2011). Yet the number of participants had little impact of the type of methodology used, as researchers were just as likely to use independent and integrated concurrent reports with large samples as with smaller ones. While some researchers relied on concurrent verbal reports (e.g. Hamada, 2009; Alsheikh, 2011), many used concurrent reports in combination with retrospective reports to facilitate qualitative analyses of verbal report data (e.g. Yang, 2006; Kim, 2011).

Studies focused on the methodology of verbal reports—Only eight studies focused on verbal reporting as a method, that is, the way in which this method was deployed to address reading processes. Among these studies, the concept of reactivity of verbal reports, particularly with reference to comprehension, appeared in six of these studies (Leow & Morgan-Short, 2004; Bowles & Leow, 2005; Goo, 2010; Jahandar *et al.*, 2012; Gurses & Adiguzel, 2013; Godfroid & Spino, 2015). In the remaining studies, concurrent and retrospective reports helped determine how participants identified target forms (Camps, 2003) and examine factors which lead to anxiety when reading in a FL (Güvendir, 2014). Overall, researchers employed concurrent and metalinguistic verbal reports simultaneously (e.g. Bowles & Leow, 2005), concurrent verbal reports independently (e.g. Goo, 2010; Godfroid & Spino, 2015), and concurrent and retrospective verbal reports in conjunction with each other (Camps, 2003).

Across these studies, there were disparities. In one study, investigation of metalinguistic and non-metalinguistic verbal reports revealed that participants in the non-metalinguistic condition performed better on comprehension measures than participants in the metalinguistic condition (Bowles & Leow, 2005). Similarly, a positive significant effect of verbal protocols on readers' comprehension was reported in Jahandar *et al.* (2012). In another study, thinking-aloud had no effect on EFL readers' comprehension (Godfroid & Spino, 2015). In other studies, where concurrent, introspective (or metalinguistic) reports were compared, participants' use of concurrent verbal reports in the absence of introspection/metalinguistic features had no detrimental effect on adult readers' comprehension (Leow & Morgan-Short, 2004), neither did it negatively impact the learning target (Goo, 2010). In all studies where the effects of verbal reports were investigated, researchers primarily deployed quantitative measures, generally in experimental designs.

Significant patterns and interpretive commentary—Overall, several patterns emerged in investigating the types of verbal report methodology used within and across mixed-method, quantitative, and qualitative studies. Independent concurrent verbal reports, integrated concurrent reports, and retrospective reports appeared most commonly. The concurrent method appeared equally in quantitative and mixed-method studies but was less frequent in qualitative studies. Based on these findings, we identified methodological concerns in relation to quantitative, qualitative, and mixed-methods approaches.

<u>Concerns using verbal report methods</u>—Given that most studies that reported a theoretical basis employed a cognitivist approach to verbal reports, a central concern involves the validity of verbal reports. Researchers either inconsistently described the conditions under which participants were required to verbalise or they failed to do so altogether. Absent this description, it was impossible to know the social circumstances that impinged on the production of verbal protocols. In a large number of cases where the verbal report strategy was used, participants practiced and subsequently performed the task independently before engaging in the experimental condition. However, researchers provided little or no information about the conditions under which these participants verbalised their thoughts. Consequently, the extent to which reports reflected the conditions for verbalising remained unknown (see Ericsson & Simon, 1984; Cohen, 1995).

In their original recommendations about the use of verbal reports, Ericsson and Simon (1984) noted the role of individual differences in participants' ability to produce think-aloud reports and cautioned researchers to guard against equating participants' general verbal ability with their ability to report verbally, a recommendation echoed by Spivey and King (1989). The extent to which participants are able to reflect and report thus influences results (see Cohen, 1995), and in the case of ELLs, recognising the complexity involved in the basic processes of producing protocols seems warranted. As explained by Ericsson and Simon (1984):

> Persons fluent in a second language can usually think aloud in that language even while thinking internally in the oral code of their native language or in non-oral code. In this case, there is nearly a one-to-one mapping between structures in the oral code of the first language and the code of the second language that is used for vocalization. How much the thinking is slowed down will then be a function of the subject's skill in the second language. (p. 250)

While researchers often cite the training provided to participants for its capacity to enhance participants' verbal reporting abilities, most studies, unlike the previous work of Jiménez *et al.* (1995), provide no evidence to indicate that participants' verbalisation capacities were assessed, nor is there any acknowledgement of the possible impact this factor might have had on findings (see Cohen, 1995). The majority of studies provided participants with instructions that purportedly allowed them to verbalise in a preferred language as well as to reduce constraints on verbalisation. However, oral- or written-language proficiencies need not be consistent with inherent capacity to report the contents of memory; therefore, assessments of language proficiency are not necessarily an accurate representation of participants' verbalisation

capacities. The absence of this distinction between linguistic capacity and deploying that capacity strategically may have inhibited the identification of differences in verbalisation and thereby affected comparisons in many studies reviewed. Jiménez *et al.* (1995) showed, when eliciting think-alouds from two bilingual Latina speakers and a monolingual English speaker while exploring cognitive and metacognitive knowledge during reading, that considering both the bilingual students' ability and willingness to think aloud while reading in Spanish (i.e. and not the target language, English) as well as 'the bilingual students' ability to think aloud while simultaneously reading silently' (p. 71) were crucial to obtaining a representative account of participants' thinking about reading. The researchers (see Jiménez *et al.*, 1996) also allowed their participants to think-aloud during their reading of both Spanish and English texts as well as to engage in both prompted and unprompted think-alouds to facilitate representativeness of reporting in ways that acknowledged linguistic difference. Such practices can prove useful for studies that deploy verbal reporting with second-language speakers and bilingual or multilingual learners.

Concerns with quantitative, qualitative and mixed-methods analyses—Verbal reports as a methodological tool, and as conceived within a cognitivist perspective, intend to capture the contents of memory associated with the reading process (as well as various other tasks, depending upon the study goals). In utilising this tool, the majority of studies relied on mixed-methods and to a lesser degree quantitative approaches. The preponderance of mixed-method approaches suggests that researchers initially used quantitative analysis and then employed qualitative techniques to further explore specific instances or cases. However, as we have illustrated, though mixed-method studies primarily reflected qualitative data collection methods, they relied heavily on quantitative analysis of the data collected, leaving a gap in the exploration of how reading processes occur.

The prominence of quantitative results from verbal protocol research echoes Fitzgerald's (1995) findings. While quantitative research is warranted in the context of LL reading studies, exploring reading processes via qualitative research through verbal reports remains crucial. In fact, capitalising on qualitative information within quantitative studies could illuminate specific reading processes while minimising the focus on the products of reading. The recognition that verbal reports may not merely be used as a means of deriving information for coding in quantitative analysis but can also serve as a functional tool for understanding the qualitative processes of readers, implies that more may be done to explore the manner in which LLs accomplish reading tasks. It is also important to recognise the use of qualitative methods within a quantitatively conceived study of literacy is not the same approach as an exploratory, interpretive study of literacy in context.

Though few in number, the instances in which researchers used a three-pronged approach to mixed-method studies—from qualitative to quantitative and then back to qualitative—attempt to present a holistic view of LLs' reading processes. This three-pronged approach presented an opportunity to rely on *qualitative* interview and observation data, and therefore allowed for access to the 'how' of L2 reading through identification of nuances in language use during verbalisation. In contrast, the opportunity to then transition to a reliance on quantitative data allowed for access to the

'what' of L2 reading through taxonomies that clearly articulated entities within reading strategies and processes as well as the connections across these elements. Furthermore, the pattern noted earlier in which integrated concurrent and retrospective reports functioned particularly well within qualitative studies seems to suggest that the combination of other methods of verbal reports with concurrent methodologies may provide greater insight into the ways in which multilingual students make sense of L1, L2 and L3 reading. Adoption of a qualitative→quantitative→qualitative paradigm in studies where concurrent reports are combined with alternative formats such as retrospection thus seems to have the potential to provide insight into what students do in reading while also clarifying reasons for students' uses of certain reading processes as well as the manner in which those processes transpire during reading. In other words, in cases where researchers collected protocol information, subsequently conducted quantitative analyses and described the underlying reasons and factors (see Chun, 2001; Gascoigne, 2002), it was more useful when retrospective protocols were also collected.

Beyond this, a more complex qualitative–quantitative–qualitative approach seems inadequate even when researchers use solely cognitivist approaches and information-processing models. As discussed earlier, in the field of applied linguistics, and considering the work of Firth and Wagner (1997) that acknowledged the need for examining language use and acquisition within the social contexts in which they are situated, it is possible to consider the socially situated nature of verbal protocols in data collection, transcription and analysis within the context of verbal reporting that focuses on literacy processes (see Deschambault, 2012). Researchers are already experimenting with sociocultural approaches to collecting and analysing verbal reports for the purposes of assessing culturally and linguistically diverse learners in areas such as cognitive interviewing and field methods (e.g. Willis & Miller, 2011). Opportunities to tap into LLs' reading processes through alternative methods such as these hold promise.

Geographical, demographic, and linguistic concerns.   Our second guiding question for this review prompted us to examine the role of verbal reports when used to explore reading processes in various contexts, with multiple languages and for different learners. We now examine geographical, demographic and linguistic patterns in the studies reviewed.

Geographical context—Approximately equal numbers of research studies drew on work in the USA (e.g. Goo, 2010; Alsheikh, 2011; Wang, 2014) and non-US territories. Of the non-US territories reported, locations included Canada (three studies: e.g. Wesche & Paribakht, 2000; Nassaji, 2003; Nassaji & Hu, 2012), Korea (two studies: Ko, 2005; Kim & Cha, 2015). Each other country had one study: Austria (Kremmel *et al.*, 2015), Belgium (Godfroid & Schmidtke, 2013), China (Li & Chun, 2015), Iran (Paribakht, 2005), Iraq (Mohammed *et al.*, 2015), Libya (Bengeleil & Paribakht, 2004), Malaysia (Ali, 2013; Radzi & Aziz, 2013), New Zealand (Tabata-Sandom, 2013), Romania (Geladari *et al.*, 2010), Taiwan (Yang, 2006; Chou, 2013), Thailand (Kasemsap & Lee, 2015), the Netherlands (Stevenson *et al.*, 2007), Turkey (Akyel & Erçetin, 2009), and Singapore (Zhang *et al.*, 2007).

In the US studies, independent concurrent reports were the predominant method used (e.g. Gascoigne, 2002; Rossomondo, 2007; Alsheikh, 2011). In non-US studies, researchers tended to use multiple concurrent methods. Among these, concurrent-retrospective reports were most common (Wesche & Paribakht, 2000; Nassaji, 2003; Bengeleil & Paribakht, 2004; Paribakht, 2005; Yang, 2006; Geladari *et al.*, 2010), though in some instances, researchers used independent concurrent methods (e.g. Ko, 2005; Stevenson *et al.*, 2007; Zhang *et al.*, 2007; Akyel & Erçetin, 2009). When integrated reporting methods were used in non-US studies, researchers tended to report qualitative results that focused primarily on the ways participants used knowledge strategies, the varieties of strategies used, and the differences in the kinds of knowledge sources that LLs relied on for reading.

Age and level of education—Of the studies that referenced participants' ages, some analysed participants 15 years old or younger (e.g. Daalen-Kapteijns *et al.*, 2001; Dressler *et al.*, 2011) and another analysed subjects between 16 and 47 (e.g. Gascoigne, 2002; Alsheikh & Mokhtari, 2011; Shin & Riazantseva, 2015). Though specific age ranges for participants in numerous studies were not provided, most studies specified participants' levels of education. Of the research studies in which level of schooling was stated, nine involved K-12 students (e.g. Zhang *et al.*, 2007; Yin, 2013), 27 involved undergraduate students (e.g. Leow, 2001; Ali, 2013; Chou, 2013), and seven examined graduate students (e.g. Upton & Lee-Thompson, 2001; Lee-Thompson, 2008; Alsheikh, 2011).

In studies where verbal reports were used to explore reading at the K-12 level, the majority of participants were in middle or high school. Two exceptions involved primary grade students who were required to produce concurrent and retrospective reports (Kim, 2011) and concurrent and introspective verbal protocols (Hu & Nassaji, 2012). All other participants provided concurrent reports while reading, and information obtained from the data was primarily quantitative. The use of older students from the K-12 span is also consistent with the maturity required for the extra attention used in protocol reporting (Spivey & King, 1989).

For undergraduate students—the education level studied most frequently—concurrent reports were most common and researchers often reported quantitative characteristics of students' reading processes (e.g. Ko, 2005; Weil, 2008; Horiba, 2013). Even in the few instances where integrated concurrent methods were used with undergraduate students (Chun, 2001; Upton & Lee-Thompson, 2001; Camps, 2003; Paribakht, 2005; Vorobel & Kim, 2011), researchers reported quantitative data such as comparisons between first and second semester students who maintained higher scores (Camps, 2003) and learner performance on tests with the presence or absence of an internal glossary (Chun, 2001). Three studies in which integrated concurrent methods were used deviated from this norm (Paribakht, 2005; Weil, 2008; Vorobel & Kim, 2011). For instance, Paribakht (2005) used concurrent introspective reports to access the variety of knowledge sources that students used when reading as well as instances of inferring lexical versus non-lexicalised target words. Similarly, Vorobel and Kim (2011) relied on concurrent retrospective reports to identify themes such as using hypermedia, dialoguing, previewing and determining what to read in participants' online reading strategies. Weil (2008) also followed a process of coding,

identifying categories from participants' protocols based on their concurrent think-alouds as they deployed strategies while reading a challenging passage.

In the few studies of graduate students' verbal reports, integrated concurrent and independent concurrent measures were used, yet we observed no clear patterns in their use. In summary, think-alouds tend to be conducted within educational contexts, in independent concurrent format and using quantitative analyses regardless of the grade levels and ages of participants.

Languages and language use—We gathered information about the languages used and the nature of language use (i.e. bilingual, multilingual, L1, L2). Certain studies also indicated participants' levels of L1 and L2 proficiency. Despite their focus on reading, very few studies reported participants' levels of reading proficiency in L2.

Researchers used verbal reports with a wide array of language types. However, languages often served very different functions. The most common language observed was English, used as both a second language and a native language. Participants who were native English speakers typically produced verbalisations in English (e.g. Gascoigne, 2002; Karnal & Pereira, 2015); in some instances where a foreign language was being learned, no reference was made to the language of verbal reporting (e.g. German: Chun, 2001; Spanish: Goo, 2010). Non-native English-speaking participants varied widely in their proficiencies with L2s as measured by a variety of comprehension and language assessments; for these learners, verbalisations were registered in either the native language (e.g. *Chinese*: Nassaji, 2003; *Farsi*: Paribakht, 2005; *Korean*: Weil, 2008) or in English (e.g. Comer, 2012a; Li & Suen, 2013). Two studies concentrated on Spanish learners' verbalisations (Leow & Morgan-Short, 2004; O'Donnell, 2009), and one study, much like the research undertaken by Jiménez *et al.* (1996) that allowed students to verbalise in both Spanish and English, required graduate students to provide concurrent reporting of thoughts as they read in three target languages, Hausa, English and French (i.e. Alsheikh, 2011). Because of inconsistencies in the number and variety of languages functioning as L1 and L2, it was difficult to determine whether patterns existed in verbal report use and to identify patterns.

Notwithstanding these difficulties, the many instances where participants produced verbal reports in English drew from both concurrent and integrated concurrent methods, and despite the verbal report method used in these studies, researchers deployed qualitative analyses, making it possible to report factors such as varieties of strategies used (Alsheikh, 2011), differences arising from vocabulary, orthography, grammar, and background knowledge (Lee-Thompson, 2008), and the ways in which ESL learners relied on their L1 and L2 (Upton & Lee-Thompson, 2001) during reading. Specifically, the findings indicated a limited use of reading strategies in native language as compared with English and French, with most proficient readers using a greater variety of strategies (Alsheikh, 2011); a total of 12 bottom-up and 15 top-down strategies used in text comprehension where common difficulties experienced by learners were vocabulary, orthography, grammar and background (Lee-Thompson, 2008); and a situation where intermediate ESL students tended to think about and process the L2 reading task using their L1 more frequently than advanced ESL students (Upton & Lee-Thompson, 2001).

Language proficiency and reading proficiency—Research studies that allowed participants to choose the verbalisation language based on how proficient they were tended to focus on intermediate learners of English (e.g. Ko, 2005; Akpinar, 2013). Fewer studies reported including advanced proficiency students (Bengeleil & Paribakht, 2004; Akyel & Erçetin, 2009; Hu & Nassaji, 2012) and low proficiency LLs (O'Donnell, 2012; Kasemsap & Lee, 2015).

A few researchers used verbal reports to explore the role of language or reading proficiency in readers' performance. Among these was Dressler *et al.*'s (2011) investigation of 12 fifth-grade Spanish-speaking ELLs' concurrent verbalisations of cognate knowledge in assigning meaning to English words. In addition, researchers used concurrent retrospective and concurrent introspective reports to determine the effects of university students' L2 lexicalisation on inferencing from the perspective of EFL learners' L2 proficiency (Bengeleil & Paribakht, 2004) and from the standpoint of L1 lexicalisation of target English words (Paribakht, 2005). Upton and Lee-Thompson (2001), also working with university-level L2 students, relied on concurrent and retrospective reports to understand the reading proficiencies of L2 learners and how they used their L2 to understand L2 general expository text. Interestingly, the findings across studies differed significantly.

Significant patterns and interpretive commentary—Researchers primarily focused on verbal reports of EFL and ESL learners, with an emphasis on intermediate-level learners. The majority of reports came from undergraduate-level students and intermediate or high school learners at the K-12 level; across these populations, concurrent reporting tended to be the norm, with a focus on quantitative results. In most studies, researchers did not acknowledge effects of individual differences in verbal reporting or provide information about the languages in which participants verbalised and the relationship with participants' developmental and linguistic backgrounds. The failure of researchers to provide participants' ages has been previously identified as problematic (e.g. Bernhardt, 2005). Beyond this, clear evidence from studies such as that conducted by Pritchard (1990) where think-alouds were elicited from adolescent American and Palauan students, indicates that the process of verbal reporting can provide information about how students' cultural background influences the way in which reading is processed and comprehended.

Although relatively few studies focused on K-12 learners, and even fewer on the lower elementary grades, learners' age differences emerged as one of the important factors. Fitzgerald (1995) noted researchers' tendency to focus on adult learners when studying reading processes using verbal reports; this tendency is most likely related to Ericsson and Simon's (1993) observation that younger learners are less likely to be able to report their thought processes. However, as Pressley and Afflerbach (1995) note, the assumption that a certain educational level automatically translates into expert reader status fails to account for individual differences among readers and variations in their knowledge, experiences, and interpretations of texts. In fact, much like third- and fourth-grade users' think-alouds in elementary foreign language immersion programmes reflected the strategies they used during classroom reading and writing (Chamot & El Dinary, 1999), the success experienced with younger students from one study in this review shows that verbal reports *can* elicit

substantive information from such students (Kim, 2011). This evidence is aligned with previous research conducted by Chamot and El Dinary (1999) that showed that the use of think-alouds with third- and fourth-grade learners in *elementary* foreign language immersion programmes helped to illustrate strategies they used to engage in reading and writing tasks.

Given that concerns about validity and reactivity are the most prevalent reasons for not using verbal reports with younger learners, a sociocultural perspective may be a fruitful alternative.

Most researchers did not attend to individual differences in participants' developmental and linguistic backgrounds. Overall, there was limited information about social contexts and learner backgrounds. There were occasional exceptions (Alsheikh, 2011). As Bernhardt (2005) notes, many studies involve LL participants from similar multilingual backgrounds, but whose experiences with any given language vary. With little information about participants' individual developmental and language learning backgrounds, it is difficult to account for variations among students. We would need to know more about year of first exposure to the L1, time spent learning the L1, number of countries in which students lived, number of languages spoken in the country in which students lived, language predominantly spoken in the home, and language officially learned by students in school. A fuller account of LLs' reading processes would require more robust data about their varied backgrounds—their home languages and cultures—and analysis of this information to inform the verbal reports produced within a given context, all of which are influenced by participants' diverse backgrounds (NCTE, 2008).

Conclusion

This methodological review synthesised original studies published between 2000 and 2015 in which verbal reports were used to gather information about the reading processes of LLs. We observed: (a) a focus on strategy use, vocabulary, comprehension, and technology when using verbal report methods with LLs; (b) a tendency to report quantitative information regardless of the type of verbal reporting method and the component of reading explored; (c) the predominance of independent concurrent methods in studies that emphasise the reading *product*, and of integrated verbal reports (i.e. concurrent and other verbal report methods) in sociocultural studies that reflect the reading *process*; (d) a heavy reliance on L1 reading and language models to undergird studies in instances where theoretical frameworks were present; (e) concerns about validity in studies based on cognitivist models of verbal reports; (f) a tendency to use concurrent verbal reports (independently or in conjunction with other verbal report methods) in studies; (g) a tendency to use integrated concurrent methods in qualitative studies; (h) a tendency to use a qualitative→quantitative versus a qualitative→quantitative→qualitative verbal reporting paradigm across studies; (i) controversial effects of concurrent (non-metalinguistic) and introspective (metalinguistic) verbal reporting methods on LLs' comprehension and acquisition of tense; (j) a reliance on solely quantitative measures to explore the ways in which verbal reports function for LLs; (k) a tendency to use independent concurrent methods and to report quantitative results in US contexts, and a tendency to use integrated

concurrent methods and to report qualitative results beyond the USA; (l) a dearth of information about the languages used by participants to verbalise thought; (m) a focus on ESL and EFL learners (as opposed to bilingual or multilingual speakers) for verbal reporting; (n) an inclination to mention the language proficiency of participants but not their *reading* proficiency; and (o) a focus on intermediate-level learners of English.

Recommendations

Based on these findings, we make several recommendations for improving the use of verbal reports to study the reading processes of LLs.

Embrace a more holistic view of verbal reports. A general consensus exists (Fitzgerald, 1995; Bowles, 2008; Cohen, 2013) that more systematic research is needed to facilitate the modification of the verbal report tool for use with ELLs. A holistic view of reading research involving verbal reports will require us to dismantle dichotomies that adopt a single perspective, and instead look towards an approach where sociocultural and cognitivist approaches are both used to explain how LLs talk about their thinking during reading. This recommendation is aligned with the reconceptualisation of research in SLA proposed by Firth and Wagner (1997), who called for an extension of examinations of language use and acquisition beyond cognitivist perspectives towards orientations that foregrounded the social and contextual factors involved in this process.

Future studies with verbal reports might productively consider Vygotsky's (1987) account of how participants' verbalisations regulate their evolving conceptualisations while reading. As a tool that elicits 'talk about thinking', verbal reports may productively be used in literacy research to study the social nature of speech (Smagorinsky, 2011). This perspective, which highlights the importance of the sociocultural in reading while also attending to the cognitive capacities of the reader, provides an alternative to solely cognitive uses of verbal reports (e.g. Bowles & Leow, 2005; Bowles, 2010a,b). A focus on whether contents of the mind can be directly accessed in verbal reports is often useful. However, attention to the negotiation that occurs within the context of the 'conversation' between the participant and researcher, a perspective more in line with a sociocultural approach, can also illuminate L2 reading (Smith & King, 2013). To investigate LLs' reading processes from a sociocultural perspective would require researchers to use more qualitative analyses of reports. Considering the social factors embedded in the reading task, as well as the interactions between researcher and participants, could be useful.

Systematically explore students' thinking in conjunction with multimodal forms of literacy, ranging from the Internet to other technological and mobile tools. Researchers would do well to explore the use of verbal report measures for studying LLs as they develop online reading comprehension characterised by 'reading to define important questions', 'reading to locate information', 'reading to evaluate information', 'reading to synthesise information' and 'reading and writing to communicate information' (Leu *et al.*, 2012). Although the nature and definition of literacy has evolved significantly

(Coiro *et al.*, 2008; Lankshear & Knobel, 2011; The New London Group, 1996; Street, 1995), we do not yet know enough about whether and how verbal reports can help analyse social practices surrounding LLs' literacy development. Park and Kim (2016) explored how five ELLs used 'hybrid' online reading strategies, transferring existing reading strategies from reading in print to reading in online contexts, but more work along these lines is needed. Researchers' emphasis on traditional forms of reading fails to capture the more dynamic processes occurring in this era of multiliteracies.

Attend to the reading theories underlying studies in which they use verbal reports.   When a theoretical framework was used, research on L2 reading remains grounded in L1 theoretical reading models, despite the concerns that cross-linguistic and social elements may not be fully captured by the use of verbal reports within such models (see Bernhardt, 2005, 2011; Fitzgerald, 1995, 2005; Grabe, 2009; Kim, 2011). Researchers should ensure that there is an appropriate theoretical account of how LLs read underlying their work. The studies relying on L2 reading and language models in this study provide a basis for doing so. Alternatively, grounded theory (Glaser & Strauss, 1967) can be used to generate data from verbal reports for constant comparative analyses. Reading and language models that take into account cross-linguistic processes involved in bi- and multilingual contexts, as well as focusing on theories that explain the reading processes of LLs, can benefit L2 and multilingual students in the field.

Consider using verbal reports when examining LLs' reading processes within the elementary grades, both in the USA and beyond.   Studies in this review illustrate that LLs in the primary grades are capable of verbalising thought while they read. As illustrated in this review, this tool holds promise for studying literacy in younger students, especially when integrated verbal reporting measures are drawn upon and a sociocultural approach is used.

Use multiple forms of verbal reporting (i.e. concurrent–introspective, concurrent–retrospective) within studies.   The preponderance of concurrent methods across the board suggest that there is a need for multiple forms of verbal reporting as a means of capturing linguistic and metalinguistic processes in reading. Moreover, the tendency to obtain quantitative information about LLs' reading processes through the use of concurrent methods suggests that there is much to be learned from integrating concurrent measures with other forms of verbal reporting.

Place more emphasis on the value of qualitative inquiry to LL and SLA research as a means of elucidating the reading process.   Rather than functioning primarily as a tool for coding categories in preparation for quantitative analysis, qualitative inquiry may provide vivid depictions of reading. In addition, qualitative inquiry allows for examination of how individual differences and learner language backgrounds influence reading. In this context, combining concurrent and retrospective reports to explore how LLs make sense of text can prove useful.

   The recommendations based on this review are in no way exhaustive but suggest several promising ways forward for using and examining verbal report measures in

the LL reading field. We trust that researchers will use these insights to continue to expand their research on verbal reports as a mechanism for better understanding LLs' reading processes, both within and beyond the USA.

NOTE

[1] *Enhanced input* in Leow (2001) constituted a written text with the following modifications: (1) shortening the text from 384 words of the unenhanced text to 242 when leaving out details, (2) 'underlining all verb forms', and (3) bolding only verb endings (the formal imperative morpheme): 'Haga y ponga una lista en la puerta de la nevera' (p. 500). The underlying goal of making these enhancements was to attract participants' attention to the target forms and to enhance their noticing or processing. The same text without these modifications was called *unenhanced*

References

Akpinar, K. D. (2013) Lexical inferencing: Perceptions and actual behaviours of Turkish English as a Foreign Language Learners' handling of unknown vocabulary, *South African Journal of Education*, 33(3), 1–17. http://sajournalofeducation.co.za/index.php/saje/article/view/755.

Akyel, A. & Erçetin, G. (2009) Hypermedia reading strategies employed by advanced learners of English, *System*, 37(1), 136–152. https://doi.org/10.1016/j.system.2008.05.002.

Ali, A. M. (2013) A comparative study on reasoning strategies in L1 and L2 critical reading-thinking tests, *World Applied Sciences Journal*, 21, 1–11. https://doi.org/10.5829/idosi.wasj.2013.21.sltl.2131.

Alsheikh, N. O. (2011) Three readers, three languages, three texts: The strategic reading of multilingual and multiliterate readers, *The Reading Matrix*, 11(1), 34–53. http://www.readingmatrix.com/articles/january_2011/alsheikh.pdf.

Alsheikh, N. O. & Mokhtari, K. (2011) An examination of the metacognitive reading strategies used by native speakers of Arabic when reading in English and Arabic, *English Language Teaching*, 4 (2), 151–160. https://doi.org/10.5539/elt.v4n2p151.

Anderson, J. R. (2005) *Cognitive psychology and its implications* (6th edn) (New York, Worth).

Bengeleil, N. F. & Paribakht, T. S. (2004) L2 reading proficiency and lexical inferencing by university EFL learners, *The Canadian Modern Language Review*, 61(2), 225–249. https://doi.org/10.3138/cmlr.61.2.225.

Bernhardt, E. B. (2005) Progress and procrastination in second language reading, *Annual Review of Applied Linguistics*, 25, 133–150. https://doi.org/10.1017/s0267190505000073.

Bernhardt, E. B. (2011) *Understanding advanced second-language reading* (New York, Routledge).

Bialystok, E. (2001) *Bilingualism in development: Language, literacy, and cognition* (New York, Cambridge University Press).

Biber, D., Nekrasova, T. & Horn, B. (2011). The effectiveness of feedback for L1-English and L2-writing development: A meta-analysis(TOEFL iBT Research Report No. TOEFL iBT-14). Available online at: http://www.ets.org/Media/Research/pdf/RR-11-05.pdf.

Bowles, M. (2004) L2 glossing: To CALL or not to CALL, *Hispania*, 87(3), 541–552. https://doi.org/10.2307/20063060.

Bowles, M. (2008) Task type and reactivity of verbal reports in SLA: A first look at a L2 task other than reading, *Studies in Second Language Acquisition*, 30(3), 359–387. https://doi.org/10.1017/s0272263108080492.

Bowles, M. (2010a) *The think-aloud controversy in second language research* (New York, Routledge).

Bowles, M. (2010b) Concurrent verbal reports in second language acquisition research, *Annual Review of Applied Linguistics*, 30, 111–127. https://doi.org/10.1017/s0267190510000036.

Bowles, M. A. & Leow, R. P. (2005) Reactivity and type of verbal report in SLA research methodology: Expanding the scope of investigation, *Studies in Second Language Acquisition*, 27, 415–440. https://doi.org/10.1017/s0272263105050187.

Brown, C. (1998) L2 reading: An update on relevant L1 research, *Foreign Language Annals*, 31, 191–202. https://doi.org/10.1111/j.1944-9720.1998.tb00567.x.

Camps, J. (2003) Concurrent and retrospective verbal reports as tools to better understand the role of attention in second language tasks, *International Journal of Applied Linguistics*, 13(2), 201–221. https://doi.org/10.1111/1473-4192.00044.

Chamot, A. U. (2004) Issues in language learning strategy research and teaching, *Electronic Journal of Foreign Language Teaching*, 1(1), 14–26. http://e-flt.nus.edu.sg/v1n12004/chamot.pdf.

Chamot, A. U. & El Dinary, P. B. (1999) Children's learning strategies in language immersion classrooms, *The Modern Language Journal*, 83(3), 319–338. https://doi.org/10.1111/0026-7902.00025.

Cheung, A. C. & Slavin, R. E. (2012) Effective reading programs for Spanish-dominant English language learners in the elementary grades: A synthesis of research, *Review of Educational Research*, 82(4), 351–395. https://doi.org/10.3102/0034654312465472.

Chou, M. H. (2013) Strategy use for reading English for general and specific academic purposes in testing and nontesting contexts, *Reading Research Quarterly*, 48(2), 175–197. https://doi.org/10.1002/rrq.42.

Chun, D. (2001) L2 reading on the web: Strategies for accessing information in hypermedia, *Computer Assisted Language Learning*, 14(5), 367–403. https://doi.org/10.1076/call.14.5.367.5775.

Cohen, A. D. (1983) Introspecting about second language learning, *Studia Anglica Posnaniensia*, 15, 149–156. https://sites.google.com/a/umn.edu/andrewdcohen/publications/research-methodology.

Cohen, A. D. (1987) Using verbal reports in research on language learning, in: C. Faerch & G. Kasper (Eds) *Introspection in second language research* (Clevedon, UK, Multilingual Matters), 82–95.

Cohen, A. D. (1995) In which language do/should multilinguals think?, *Language, Culture, and Curriculum*, 8(2), 99–113. https://doi.org/10.1080/07908319509525194.

Cohen, A. D. (1996) Verbal reports as a source of insight into second language learner strategies, *Applied Language Learning*, 7(1–2), 5–24. https://sites.google.com/a/umn.edu/andrewdcohen/publications/research-methodology.

Cohen, A. D. (2013) Verbal report, in: C. A. Chapelle (Ed.) *The encyclopedia of applied linguistics* (Oxford, UK, Wiley-Blackwell), 1–5.

Coiro, J., Knobel, M., Lankshear, C. & Leu, D. (2008) *The handbook of research on new literacies* (Mahwah, NJ, Erlbaum).

Comer, W. J. (2012a) Lexical inferencing in reading L2 Russian, *Reading in a Foreign Language*, 24(2), 209–230. http://nflrc.hawaii.edu/rfl.

Comer, W. J. (2012b) The role of grammatical knowledge in reading for meaning in Russian, *The Slavic and East European Journal*, 56(2), 232–255. http://www.jstor.org/stable/41698474.

Comer, W. J. (2014) Reading L2 Russian: The challenges of the Russian–English dictionary, *The Reading Matrix*, 14(2), 1–19. http://www.readingmatrix.com/files/11-tx75m475.pdf.

Common Core State Standards Initiative. (2010) Implementing the Common Core State Standards. Available online at: http://www.corestandards.org/

Daalen-Kapteijns, M., Elshout-Mohr, M. & De Glopper, K. (2001) Deriving the meaning of unknown words from multiple contexts, *Language Learning*, 51(1), 145–181. https://doi.org/10.1111/0023-8333.00150.

Deschambault, R. (2012) Thinking-aloud as talking-in-interaction: Reinterpreting how L2 lexical inferencing gets done, *Language Learning*, 62(1), 266–301. https://doi.org/10.1111/j.1467-9922.2011.00653.x.

Dixon, L. Q., Zhao, J., Shin, J.-Y., Wu, S., Burgess-Bringham, R., Gezer, M. U. & Snow, C. (2012) What we know about second language acquisition: A synthesis from four perspectives, *Review of Educational Research*, 82(1), 5–60. https://doi.org/10.3102/0034654311433587.

Dressler, C., Carlo, M. S., Snow, C. E., August, D. & White, C. E. (2011) Spanish-speaking students' use of cognate knowledge to infer the meaning of English words, *Bilingualism: Language and Cognition*, 14(2), 243–255. https://doi.org/10.1017/s1366728910000519.

Ericsson, K. A. & Simon, H. A. (1984/1993). *Protocol analysis: Verbal reports as data* (Cambridge, MA, MIT Press).

European Commission. (2012) Foreign language learning statistics Available online at: http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Foreign_language_learning_statistics.

Firth, A. & Wagner, J. (1997) On discourse, communication, and (some) fundamental concepts in SLA research, *The Modern Language Journal*, 81(3), 285–300. https://doi.org/10.2307/329302.

Firth, A. & Wagner, J. (1998) SLA property: No trespassing!, *The Modern Language Journal*, 82(1), 91–94. http://www.jstor.org/stable/328686.

Fitzgerald, J. (1995) English-as-a-second-language learners' cognitive reading processes: A review of research in the United States, *Review of Educational Research*, 65(2), 145–190. https://doi.org/10.3102/00346543065002145.

Fitzgerald, J. (2005) Multilingual writing in preschool through twelfth grade: The last 15 years, in: C. A. MacArthur, S. Graham & J. Fitzgerald (Eds) *Handbook of writing research* (New York, Guilford Press), 337–354.

Gascoigne, C. (2002) Documenting the initial language reading experience: The readers speak, *Foreign Language Annals*, 35(5), 554–560. https://doi.org/10.1111/j.1944-9720.2002.tb02723.x.

Gass, S. M. & Mackey, A. (2007) Input, interaction, and output in second language acquisition, in: B. Van Patten & J. Williams (Eds) *Theories in second language acquisition: An introduction* (Mahwah, NJ, Lawrence Erlbaum), 175–199.

Geladari, A., Griva, E. & Mastrothanasis, K. (2010) A record of bilingual elementary students' reading strategies in Greek as a second language, *Procedia—Social and Behavioral Sciences*, 2(2), 3764–3769. https://doi.org/10.1016/j.sbspro.2010.03.585.

Ghavamnia, M., Ketabi, S. & Tavakoli, M. (2013) L2 reading strategies used by Iranian EFL learners: A think-aloud study, *Reading Psychology*, 34(4), 355–378. https://doi.org/10.1080/02702711.2011.640097.

Ghonsooly, B. & Barghchi, B. (2012) An introspective study of L2 reading anxiety, *Journal of Teaching Language Skills*, 3(1), 65–98. http://jtls.shirazu.ac.ir/?_action=articleInfo&article=394.

Glaser, B. & Strauss, A. L. (1967) *The discovery of grounded theory: Strategies for qualitative research* (Chicago, IL, Aldine).

Godfroid, A. & Schmidtke, J. (2013) What do eye movements tell us about awareness? A triangulation of eye-movement data, verbal reports, and vocabulary learning scores, in: J. M. Bergsleithner, S. N. Frota & J. K. Yoshioka (Eds) *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (Honolulu, University of Hawai'i, National Foreign Language Resource Center), 183–205.

Godfroid, A. & Spino, L. A. (2015) Reconceptualizing reactivity of think-alouds and eye tracking: Absence of evidence is not evidence of absence, *Language Learning*, 65(4), 896–928. https://doi.org/10.1111/lang.12136.

Goo, J. (2010) Working memory and reactivity, *Language Learning*, 60(4), 712–752. https://doi.org/10.1111/j.1467-9922.2010.00573.x.

Goodman, K. S. (1967) Reading: A psycholinguistic guessing game, *Journal of the Reading Specialist*, 6, 126–135. https://doi.org/10.1080/19388076709556976.

Goodman, K. S. (1968) *The psycholinguistic nature of the reading process* (Detroit, MI, Wayne State University Press).

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

Gurses, M. O. & Adiguzel, O. C. (2013) The effect of strategy instruction based on the Cognitive Academic Language Learning Approach over reading comprehension and strategy use, *Journal of Education and Learning*, 2(2), 55–68. https://doi.org/10.5539/jel.v2n2p55.

Güvendir, E. (2014) Using think-aloud protocols to identify factors that cause foreign language reading anxiety, *The Reading Matrix*, 14(2), 109–118. http://www.readingmatrix.com/files/11-bkw8c881.pdf.

Hamada, M. (2009) Development of L2 word-meaning inference while reading, *System*, 37(3), 447–460. https://doi.org/10.1016/j.system.2009.03.003.

Hamada, M. & Park, C. (2011) Word-meaning inference: A longitudinal investigation of inference accuracy and strategy use, *Asian EFL Journal*, 13(4), 10–32. http://asian-efl-journal.com/.

Han, Z. & Anderson, N. (2009) *Second language reading research and instruction: Crossing the boundaries* (Ann Arbor, MI, Michigan University Press).

He, T. (2008) Reading for different goals: The interplay of EFL college students' multiple goals, reading strategy use and reading comprehension, *Journal of Research in Reading*, 31(2), 224–242. https://doi.org/10.1111/j.1467-9817.2007.00355.x.

Hijikata, Y., Nakatani, Y. & Shimizu, M. (2013) Japanese EFL students' reading processes for academic papers in English, *Journal of Education and Learning*, 2(1), 70–83. https://doi.org/10.5539/jel.v2n1p70.

Horiba, Y. (2013) Task-induced strategic processing in L2 text comprehension, *Reading in a Foreign Language*, 25(2), 98–125. http://nflrc.hawaii.edu/rfl.

Hu, H. C. M. & Nassaji, H. (2012) Ease of inferencing, learner inferential strategies, and their relationship with the retention of word meanings inferred from context, *Canadian Modern Language Review*, 68(1), 54–77. https://doi.org/10.3138/cmlr.68.1.054.

Hu, H. C. M. & Nassaji, H. (2014) Lexical inferencing strategies: The case of successful versus less successful inferencers, *System*, 45, 27–38. https://doi.org/10.1016/j.system.2014.04.004.

Huber, C. R. & Kuncel, N. R. (2016) Does college teach critical thinking? A meta-analysis, *Review of Educational Research*, 86(2), 431–468. https://doi.org/10.3102/0034654315605917.

Jahandar, S., Khodabandehlou, M., Seyedi, G. & Abadi, R. M. D. (2012) The think-aloud method in EFL reading comprehension, *International Journal of Scientific & Engineering Research*, 3(9), 1–9. http://www.ijser.org/researchpaper%5CThe-Think-aloud-Method-in-EFL-Reading-Comprehension.pdf.

Jiménez, R. T., García, G. E. & Pearson, P. D. (1995) Three children, two languages, and strategic reading: Case studies in bilingual/monolingual reading, *American Educational Research Journal*, 32(1), 67–97. https://doi.org/10.3102/00028312032001067.

Jiménez, R. T., García, G. E. & Pearson, P. D. (1996) The reading strategies of bilingual Latina/o students who are successful English readers—opportunities and obstacles, *Reading Research Quarterly*, 31(1), 90–112. http://www.jstor.org/stable/748241.

Karimi, M. N. & Alibakhshi, G. (2014) EFL learners' text processing strategies across comprehension vs. integration reading task conditions, *System*, 46, 96–104. https://doi.org/10.1016/j.system.2014.07.013.

Karnal, A. R. & Pereira, V. V. (2015) Reading strategies in a L2: A study on machine translation, *The Reading Matrix*, 15(2), 69–79. http://www.readingmatrix.com/files/13-5728575a.pdf.

Kasemsap, B. & Lee, H. Y. H. (2015) L2 reading in Thailand: Vocational college students' application of reading strategies to their reading of English texts, *The Reading Matrix*, 15(2), 101–117. http://www.readingmatrix.com/files/13-p0gvv09k.pdf.

Kern, R. G. (1994) The role of mental translation in second language reading, *Studies in Second Language Acquisition*, 16(4), 441–461. https://doi.org/10.1017/s0272263100013450.

Ketabi, S., Ghavamnia, M. & Rezazadeh, M. (2012) Hypermedia reading strategies used by Persian graduate students in TEFL—a think-aloud study, *The Reading Matrix*, 12(1), 39–49. http://www.readingmatrix.com/articles/April_2012/ketabi_ghavamnia_rezazadeh.pdf.

Kim, D. (2011) Dialogic meaning construction and reading domains among four young English learners in second–language reading, *Multilingual Education*, 1, 1–21. https://doi.org/10.1186/2191-5059-1-2.

Kim, D. (2018) Elementary English learners' use of reading strategies with culturally relevant and culturally distant stories, *Journal of Language, Identity and Education*, 18(2), 1–19. https://doi.org/10.1080/15348458.2018.1503960.

Kim, H. I. & Cha, K. A. (2015) Korean learners' metacognition in reading using think-aloud procedures with a focus on regulation of cognition, *English Language Teaching*, 8(6), 178–193. https://doi.org/10.5539/elt.v8n6p178.

Ko, M. H. (2005) Glosses, comprehension and strategy use, *Reading in a Foreign Language*, 17 (2).125–143. http://nflrc.hawaii.edu/rfl/October2005/ko/ko.html.

Koda, K. (2007) Reading and language learning: Crosslinguistic constraints on second language reading development, *Language Learning*, 57, 1–44. https://doi.org/10.1111/0023-8333. 101997010-i1.

Kremmel, B., Brunfaut, T. & Alderson, J. C. (2015) Exploring the role of phraseological knowledge in foreign language reading, *Applied Linguistics*, ???, 1–24. https://doi.org/10.1093/applin/amv070.

Lankshear, C. & Knobel, M. (2011) *New literacies: Everyday practices and social learning: Everyday practices and classroom learning* (New York, Open University Press).

Lee-Thompson, L. (2008) An investigation of reading strategies applied by American learners of Chinese as a foreign language, *Foreign Language Annals*, 41(4), 702–721. https://doi.org/10.1111/j.1944-9720.2008.tb03326.x.

Leow, R. P. (2001) Do learners notice enhanced forms while interacting with the L2?: An online and offline study of the role of written input enhancement in L2 reading, *Hispania*, 84(3), 496–509. https://doi.org/10.2307/3657810.

Leow, R. P. & Morgan-Short, K. (2004) To think-aloud or not to think aloud: The issue of reactivity in SLA research methodology, *Studies in Second Language Acquisition*, 26(1), 35–57. https://doi.org/10.1017/s0272263104026129.

Leow, R. P., Hsieh, H. C. & Moreno, N. (2008) Attention to form and meaning revisited, *Language Learning*, 58(3), 665–695. https://doi.org/10.1111/j.1467-9922.2008.00453.x.

Leu, D. J. Jr, Kinzer, C. K., Coiro, J. & Cammack, D. (2004) Toward a theory of new literacies emerging from the Internet and other ICT, in: R. B. Ruddell & N. Unrau (Eds) *Theoretical models and processes of reading* (Newark, DE, International Reading Association), 1568–1611.

Leu, D. J., Coiro, J., Kulikowich, J. M., Sedransk, N., Everett-Cavcopardo, H., McVerry, G. … Burlingame, C. (2012) *An initial study of online reading and Internet research ability in rich and poor school districts: Do the rich get richer and the poor get poorer?* (Vancouver, CA, Educational Research Association).

Li, J. & Chun, C. K. W. (2015) Metacognitive knowledge of tertiary-level EFL reading: Case studies in an acquisition-poor context, *Chinese Journal of Applied Linguistics*, 38(2), 166–187. https://doi.org/10.1515/cjal-2015-0010.

Li, H. & Suen, H. K. (2013) Detecting native language group differences at the subskills level of reading: A differential skill functioning approach, *Language Testing*, 30(2), 273–298. https://doi.org/10.1177/0265532212459031.

Li, H. & Suen, H. K. (2015) How do Chinese ESL learners recognize English words during a reading test? A comparison with Romance-language-speaking ESL learners, *International Multilingual Research Journal*, 9(2), 93–107. https://doi.org/10.1080/19313152.2014.995013.

Lin, W.-C., Huang, H.-T. & Liou, H.-C. (2013) The effects of text-based SCMC on SLA: A meta analysis, *Language Learning & Technology*, 17(2), 123–142. http://llt.msu.edu/issues/june2013/linetal.pdf.

Lomicka, L. (1998) To gloss or not to gloss: An investigation of reading comprehension online, *Language Learning & Technology*, 1(2), 41–50. http://llt.msu.edu/vol1num2/pdf/article2.pdf.

Mahmoudi, E. (2014) Comprehension monitoring characteristics and level of proficiency: A study of Iranian EFL learners, *International Journal of English Language Teaching*, 1(2), 64–72. https://doi.org/10.5430/ijelt.v1n2p64.

Matsuda, P. K., Canagarajah, A. S., Harklau, L., Hyland, K. & Warschauer, M. (2003) Changing currents in second language writing research: A colloquium, *Journal of Second Language Writing*, 12(2), 151–179. https://doi.org/10.1016/s1060-3743(03)00016-x.

Merriam, S. B. (2009) *Qualitative research: A guide to design and implementation* (San Francisco, CA, Jossey-Bass).

Mohammed, H. A., Majid, N. A. & Abdullah, T. (2015) The differential effect of attentional condition on subsequent vocabulary development, *English Language Teaching*, 9(1), 88–100. https://doi.org/10.5539/elt.v9n1p88.

Morgan-Short, K., Heil, J., Botero-Moriarty, A. & Ebert, S. (2012) Allocation of attention to second language form and meaning, *Studies in Second Language Acquisition*, 34(04), 659–685. https://doi.org/10.1017/s027226311200037x.

Nassaji, H. (2003) L2 vocabulary learning from context: Strategies, knowledge sources, and their relationship with success in L2 lexical inferencing, *TESOL Quarterly*, 37(4), 645–670. https://doi.org/10.2307/3588216.

Nassaji, H. & Hu, H. C. M. (2012) The relationship between task-induced involvement load and learning new words from context, *International Review of Applied Linguistics in Language Teaching*, 50(1), 69–86. https://doi.org/10.1515/iral-2012-0003.

National Center for Educational Statistics (2018) English Language Learners in public schools, Available online at: https://nces.ed.gov/programs/coe/indicator_cgf.asp.

National Clearinghouse for English Language Acquisition. (2011) FAQ: How many school-aged Limited English Proficient (LEP) students are there in the U.S? Available online at: http://www.ncela.gwu.edu/faqs/

National Council of Teachers of English. (2008) English language learners: A research policy brief produced by the National Council of Teachers of English. Available online at: http://www.ncte.org/library/NCTEFiles/Resources/Positions/Chron0308PolicyBrief.pdf

National Council of Teachers of English. (2011) 2011 NCTE Resolutions. Available online at: http://www.ncte.org/press/2011resolutions

New London Group. (1996) A pedagogy of multiliteracies: Designing social futures, *Harvard Educational Review*, 66(1), 60–92. https://doi.org/10.17763/haer.66.1.17370n67v22j160u.

Norris, J. M. & Ortega, L. (Eds) (2006) *Synthesizing research on language learning and teaching* (Amsterdam, PA, John Benjamins).

O'Donnell, M. E. (2009) Finding middle ground in second language reading: Pedagogic modifications that increase comprehensibility and vocabulary acquisition while preserving authentic text features, *The Modern Language Journal*, 93(iv), 512–533. https://doi.org/10.1111/j.1540-4781.2009.00928.x.

O'Donnell, M. E. (2012) Second language learners' use of marginal glosses, *Foreign Language Annals*, 45(4), 543–563. https://doi.org/10.1111/j.1944-9720.2013.12004.x.

O'Malley, J. M. & O'Malley, M. J. (1990) *Learning strategies in second language acquisition* (New York, Cambridge University Press).

Paribakht, T. S. (2005) The influence of first language lexicalization on second language lexical inferencing: A study of Farsi-speaking learners of English as a foreign language, *Language Learning*, 55(4), 701–748. https://doi.org/10.1111/j.0023-8333.2005.00321.x.

Paris, S. G. (2005) Reinterpreting the development of reading skills, *Reading Research Quarterly*, 40(2), 184–202. https://doi.org/10.1598/rrq.40.2.3.

Park, H.-R. & Kim, D. (2016) English language learners' strategies for reading Computer-based texts at home and in school, *CALICO Journal*, 33(3), 380–409. https://doi.org/10.1558/cj.v33i3.26552.

Park, J., Yang, J. S. & Hsieh, Y. C. (2014) University level second language readers' online reading and comprehension strategies, *Language Learning and Technology*, 18(3), 148–172. http://llt.msu.edu/issues/october2014/parkyanghsieh.pdf.

Patton, M. Q. (2002) *Qualitative research and evaluation methods* (Thousand Oaks, CA, Sage).

People for Education. (2013) Language Support. Available online at: http://www.peopleforeducation.ca/wp-content/uploads/2013/07/language-support-2013.pdf

Peregoy, S. F. & Boyle O. F. (Eds) (2008) *Reading, writing and learning: A resource book for K-12 teacher* (5th edn) (Boston (Pearson), MA).

Pressley, M. & Afflerbach, P. (1995) *Verbal protocols of reading: The nature of constructively responsive reading* (Hillsdale, NJ, Lawrence Erlbaum Associates).

Pritchard, R. (1990) The effects of cultural schemata on reading processing strategies, *Reading Research Quarterly*, 25, 273–295. https://doi.org/10.2307/747692.

Radzi, A. H. M. & Aziz, N. H. A. (2013) Of a village Bomoh and the lottery: Content schemata influence on second language reading, *Advances in Language and Literary Studies*, 4(1), 42–50. https://doi.org/10.7575/aiac.alls.v.4n.1p.42.

Risko, V., Reid, L., Hoffman, J., Rodesiler, L., Skerrett, A., Teale, W. & Shoffner, M. (2017) *Literacy teacher preparation [Research advisory]*. International Literacy Association and National Council of Teachers of English (Newark, DE, Authors).

Rossomondo, A. E. (2007) The role of lexical temporal indicators and text interaction format in the incidental acquisition of the Spanish future tense, *Studies in Second Language Acquisition*, 29, 39–66. https://doi.org/10.1017/s0272263107070027.

Rott, S. (2005) Processing glosses: A qualitative exploration of how form-meaning connections are established and strengthened, *Reading in a Foreign Language*, 17(2), 95–124. http://www.nflrc.hawaii.edu/rfl/October2005/rott/rott.html.

Rott, S. & Gavin, B. (2015) Comprehending and learning from Internet sources: A conceptual replication study of Goldman, Braasch, Wiley, Greasser and Brodowinska (2012), *CALICO Journal*, 32(2), 323–354. https://doi.org/10.1558/cj.v32i2.25139.

Ryan, G. & Bernard, H. R. (2003) Techniques to identify themes, *Field Methods*, 15(1), 85–109. https://doi.org/10.1177/1525822x02239569.

Seng, G. H. (2007) The effects of think-aloud in a collaborative environment to improve comprehension of L2 texts, *The Reading Matrix*, 7(2), 29–45. http://www.readingmatrix.com/articles/seng/article.pdf.

Shin, C. & Riazantseva, A. (2015) Reader identity: A case study of Korean graduate students' meaning construction of an L2 literary text, *Language and Intercultural Communication*, 15(4), 600–615. https://doi.org/10.1080/14708477.2015.1061535.

Silva, T. & Brice, C. (2004) Research in teaching writing, *Annual Review of Applied Linguistics*, 24, 70–106. https://doi.org/10.1017/s0267190504000042.

Smagorinsky, P. (2011) Thinking and speech and verbal data, *Vygotsky and Literacy Research*, 2(3), 233–244. https://doi.org/10.1007/978-94-6091-696-0_9.

Smith, F. (1971) *Understanding reading* (New York, Holt, Rinehart and Winston).

Smith, P. & King, J. R. (2013) An examination of veridicality in verbal protocols of language learners, *Theory and Practice in Language Studies*, 3(5), 709–720.

Smith, B. & Lafford, B. (2009) The evaluation of scholarly activity in computer-assisted language learning, *The Modern Language Journal*, 93, 868–883. https://doi.org/10.1111/j.1540-4781.2009.00978.x.

Snow, C., Burns, M. S. & Griffin, P. (1998). Preventing reading difficulties in young children. Available online at: http://www.nap.edu/html/reading/

Spivey, N. N. & King, J. R. (1989) Readers as writers composing from sources, *Reading Research Quarterly*, 24(1), 7–26. http://www.jstor.org/stable/748008.

Stanovich, K. (2000) *Progress in understanding reading* (New York, NY, Guilford Press).

Stevenson, M., Schoonen, R. & de Glopper, K. (2007) Inhibition of compensation? A multidimensional comparison of reading processes in Dutch and English, *Language Learning*, 57(1), 115–154. https://doi.org/10.1046/j.1467-9922.2003.00241.x.

Street, B. (1995) *Social literacies* (London, UK, Longman).

Tabataba'ian M. S., & Zabihi, R. (2011) Strategies used by four Iranian EFL learners in reading ESP and GPE texts: A think-aloud case study, *World Journal of English Language*, 1(1), 53–62. https://doi.org/10.5430/wjel.v1n1p53.

Tabata-Sandom, M. (2013) The reader–text–writer interaction: L2 Japanese learners' response toward graded readers, *Reading in a Foreign Language*, 25(2), 264–282. http://nflrc.hawaii.edu/rfl/October2013/articles/tabata.pdf.

Teachers of English to Speakers of Other Languages (TESOL). (2010) Position paper on language and literacy development for young English language learners (ages 3–8). Available online at: http://www.tesol.org.

Upton, T. A. & Lee-Thompson, L. (2001) The role of the first language in second language reading, *Studies in Second Language Acquisition*, 23, 469–495. http://journals.cambridge.org/action/displayAbstract?fromPage=online&axml:id=93373.

Vorobel, O. & Kim, D. (2011) Upper-intermediate students' summarizing process in L2, *TESOL Journal*, 2(3), 330–354. https://doi.org/10.5054/tj.2011.259958.

Vorobel, O. & Kim, D. (2017) Adolescent ELLs' collaborative writing practices in face-to-face and online contexts: From perceptions to actions, *System*, 65, 78–89.

Vygotsky, L. S. (1987) The genetic roots of thinking and speech, in: R. W. Reiber & A. S. Carton (Eds) *The collected works of L.S. Vygotsky (Vol. 1)*. *Problems of general psychology* (New York, Plenum Press), 101–120.

Wang, J. (2014) Strategies for reading Chinese texts with and without pop-up dictionary for beginning learners of Chinese, *CALICO Journal*, 31(2), 244–260. https://doi.org/10.11139/cj.31.2.244-260.

Weil, N. (2008) Vocabulary size, background characteristics, and reading skill of Korean intensive English students, *Asian EFL Journal*, 10(4), 26–59. http://asian-efl-journal.com/.

Wesche, M. B. & Paribakht, T. S. (2000) Reading-based exercises in second language vocabulary learning: An introspective study, *The Modern Language Journal*, 84(2), 196–213. https://doi.org/10.1111/0026-7902.00062.

Willis, G. B. & Miller, K. (2011) Cross-cultural cognitive interviewing: Seeking comparability and enhancing understanding, *Field Methods*, 23(4), 331–341. https://doi.org/10.1177/1525822x11416092.

Yang, Y. (2006) Reading strategies or comprehension monitoring strategies?, *Reading Psychology*, 27, 313–343. https://doi.org/10.1080/02702710600846852.

Yanguas, I. (2009) Multimedia glosses and their effectiveness on L2 text comprehension and vocabulary learning, *Language Learning and Technology*, 13(2), 48–67. http://llt.msu.edu/vol13num2/yanguas.pdf.

Yilmaz, F. (2015) Turkish as a foreign language learners' awareness and use of morphology in guessing the meanings of unknown words from context: A case study, *Educational Research and Reviews*, 10(5), 675–683. https://doi.org/10.5897/err2014.2066.

Yin, Z. (2013) The more engaged, the better? An exploration on Chinese EFL learners' lexical inferencing procedure, *International Journal of Applied Linguistics and English Literature*, 2(5), 222–228. https://doi.org/10.7575/aiac.ijalel.v.2n.5p.222.

Zhang, L., Gu, P. Y. & Hu, G. (2007) A cognitive perspective on Singaporean primary school pupils' use of reading strategies in learning to read in English, *British Journal of Educational Psychology*, 78(2), 245–271. https://doi.org/10.1348/000709907x218179.