Background

Reading Models and Theories

In this section, we briefly review influential models and theories of reading. In doing so, we discuss the individual difference (ID) variables that are the most relevant within each theory and model.

Automaticity and Verbal Efficiency Theory

Automaticity is a concept with a long-standing research history in psychology. Acquiring automaticity is critical for fluent performance in various cognitive processes including reading (see Moors & De Houwer, 2006 for a review of different theoretical and empirical approaches). Although models and theories of reading reviewed here may vary in their claims and focuses, they share one common underlying assumption: in order for proficient reading to occur, all involved processes that are subject to automatization need to be automatized. Automatized processes are qualitatively different from simply speedy processes in that they are ballistic (unstoppable once it begins), reliably fast (indicated by a positive correlation between decreasing mean task response times and the Coefficients of Variance [Segalowitz & Segalowitz, 1993]), are executed without conscious control, allow parallel processing of multiple processes, and most importantly, demand little to no cognitive resources that are limited in capacity, such as working memory (Grabe & Stoller, 2019). Achieving automaticity in all reading-related processes that are subject to automatization maximizes residual working memory, which can then be used for comprehension processes that are less prone to automatization such as setting reading goals, using reading strategies, monitoring comprehension, or making inferences. The efficient reallocation of cognitive resources leads to better reading outcomes.

Perfetti’s Verbal Efficiency Theory (Perfetti & Lesgold, 1979), another influential theory of reading, also underscores the importance of automaticity. The theory prioritizes efficient word recognition processes in L2 reading comprehension. It claims that skilled readers have automatized lower-level processes necessary for word recognition. The automatization of lower-level processes, along with high lexical quality (i.e., stable and complete lexical representation), enables rapid and reliably efficient word recognition. As a result, skilled readers can allocate residual cognitive resources (e.g., working memory) for other comprehension processes, achieving successful comprehension.
Simple View of Reading

Similar to the verbal efficiency theory, the simple view of reading (Hoover & Gough, 1990; henceforth SVR) emphasizes the importance of decoding efficiency and language comprehension to achieve successful reading comprehension. Hoover and Gough (1990) defined “skilled decoding” as “efficient word recognition”. (p. 130). Decoding (or word recognition) is a process during which the reader converts letters (graphemes) to corresponding sounds (phonemes), identifies increasingly larger units such as syllables and morphemes, unpacks syntactic and semantic information from the written input, and retrieves relevant information of the word at hand from their mental lexicon. While being an essential component of many reading models, decoding was not explicitly included as a separate component in earlier models of foreign language aptitude. However, phonological and orthographic skills, both of which are fundamental to decoding, are inherently interconnected with phonemic coding abilities, one of the three components of Carroll’s (1981) modern language aptitude test (MLAT): grammatical sensitivity, phonetic coding ability, and memory capacity. Later, in the linguistic coding differences hypothesis (Sparks & Ganschow, 1991, 2001), first language (L1) and second language (L2) phonological and orthographic skills were acknowledged as a component of language aptitude.

With its drastically parsimonious formula—decoding x language comprehension = reading comprehension—the SVR predicted that reading comprehension is the product of decoding and language comprehension. In this model, the values of decoding and language comprehension abilities range from 0 (no ability) to 1 (perfect ability). Therefore, this model predicts that, in the absence of either decoding ability or language comprehension ability, reading comprehension cannot happen. The relative contributions of decoding and language comprehension are likely to vary across the L1 and L2 reader groups at different developmental stages. For example, young, novice L1 readers who are relatively new to decoding are likely to vary more in decoding ability rather than language comprehension. Once readers reach the ceiling in decoding abilities, however, language comprehension is likely to emerge as a key contributor to reading given increasingly demanding texts to comprehend. As for L2 readers, both decoding efficiency and language comprehension are likely to remain significant contributors to reading proficiency for a longer period because decoding automaticity is much more difficult to achieve in the L2 than in the L1. In addition, individual variability in knowledge sources that support linguistic comprehension in the L2 is also likely to stay large among L2 readers. Regarding the implications on ID variables, we can conjecture that due to the delayed automatization of decoding skills, the efficient allocation of cognitive resources such as working memory will also be delayed among L2 readers.

Top–down, Bottom–up, and Interactive Models of Reading

The early controversy of top–down vs. bottom–up models of reading is well–documented in Chall’s (1967) review of the literature published between 1910 and 1965. The top–down model of reading, also known as the psychological guessing game model or the whole language approach (Goodman, 1967), claimed that the reader uses linguistic and contextual knowledge to hypothesize what the upcoming text contains, even including the meaning of unknown words. Hence, this model views reading as a constant cycle of hypothesis formulation and testing via text sampling, and confirming/rejecting proposed hypotheses, i.e., a psychological guessing game where the higher–level processes direct and drive the lower processes including word recognition. In direct contrast, the bottom–up model of reading (e.g., Gough, 1972) claimed that reading comprehension involves a serial operation of lower–level processes. In this model, the reader processes graphemes, maps them onto phonemes, recognizes words, identifies structural constituents of recognized words, and builds semantic propositions serially. Another clear difference between the top–down and bottom–up
model is that while the former assumes that lower-level processes such as word recognition can be influenced by higher-level processes, the latter assumes word recognition to be a modular process that is impenetrable by higher-level processes or background knowledge.

While being polar opposites of each other in their claims, top–down and bottom–up models of reading each had serious theoretical and empirical limitations. First, the claims of the top–down model of reading were considered to be highly implausible because the speed of lower-level processes is much higher than that of hypothesis testing, which is a higher-level process. Therefore, fluent and efficient reading would be impossible if higher-level processes directed lower-level processes. Unsurprisingly, the top–down model failed to gain empirical support as eye-tracking research showed that eye movements during reading are automatic and largely constrained by word frequency (e.g., Rayner et al., 2005; Rayner & Well, 1996). On the other hand, the biggest problem of the bottom–up model was its rigidity which did not allow for the parallel operation of multiple processes and knowledge sources, which contradicts the nature of automatized processes (Moors & De Houwer, 2006). Consequently, the bottom–up model also failed to gain theoretical or empirical support, leaving reading researchers to seek alternative explanations.

In an attempt to find an alternative reading model, interactive models were proposed. While varying in detail, these models were similar in that lower-level and higher-level processes interact depending on the needs of the reader. For example, the interactive-compensatory model (Stanovich, 1980) argued that 1) unlike the top–down model, “semantic processes constrain the alternatives of lower levels but are themselves constrained by lower-level analyses” (p. 35); and 2) “processes at any level can compensate for deficiencies of any other level” (p. 36). To illustrate these claims, although contextual knowledge of the reader will help suppress the activation of context-irrelevant meanings of a word (e.g., the meaning of “a flying vehicle” will be suppressed when a reader sees “plane” in a geometry textbook), the word recognition process itself solely relies on lower-level processes. Furthermore, a reader with poor word recognition skills is likely to rely on higher-level processes such as inferencing or contextual knowledge to aid lexical access. Just like the complex interactions of top–down and bottom–up processes, interactive models of reading suggest interactions of multiple ID factors working in tandem during reading, including metacognition (a self-regulatory capacity of higher-level processes) and word recognition (a critical lower-level process built upon ID factors involving phonological coding and working memory).

**Individual Difference Variables in L2 Reading**

**Cognitive Variables**

Phonological awareness and working memory, two components of foreign language aptitude (e.g., MLAT, high-level language aptitude battery [Hi-LAB], and profile system), are essential predictors of L2 reading achievement. Phonological awareness, “a general appreciation of the sounds of speech as distinct from their meaning” (Snow et al., 1998), is a type of metalinguistic awareness used for identifying, segmenting, and manipulating sound structures such as phonemes and syllables (Grabe & Stoller, 2019). Unlike decoding, which entails processing the written language, phonological awareness develops during the pre-literate period, as early as between the sixth and twelfth month of age (Kuhl et al., 2006) in the L1. It is first used to process oral language and then, as the child learns to read, the written language. L2 readers, on the other hand, face the task of learning new phonological rules of L2 along with the rules of phoneme–grapheme correspondence. The interest in phonological awareness as a potential contributor to L2 reading is most noted in the linguistic coding differences hypothesis (Sparks & Ganschow, 1991, 2001). In brief, this hypothesis predicts that individuals experiencing difficulties in phonemic awareness (the ability to segment a stream of sounds into individual phonemes) and phonological decoding (the ability
to convert graphemes to phonemes and to integrate them to recognize words) in L1 will likely run into similar difficulties in L2. Because automatic word recognition is a necessary condition for successful reading, high levels of phonological awareness are likely an essential contributor to successful L2 reading.

Working memory is another cognitive variable that plays a central role in reading comprehension (Just & Carpenter, 1992; Wen & Skehan, 2011). Originally introduced by Baddeley and Hitch (1974), working memory is a cognitive system with the following characteristics: it is 1) multicomponential; 2) limited in capacity; 3) used for various cognitive tasks such as a computation, reasoning, or language comprehension; and 4) has two key functions—storage and processing. During reading, relevant information—such as decoded graphemes, the meaning of newly recognized words, or the proposition of the previous sentence—remains activated for a short period of time in working memory as more incoming information (e.g., remaining phonemes of the word, remaining words of the constituent, and transition to a new proposition) is supplied to form a coherent representation of the text. Working memory is also involved in higher-level processes such as online inferencing, which is necessary for comprehension.

Due to its far-reaching implications on various processes involved in language comprehension, working memory has been increasingly recognized as an essential component of foreign language aptitude (Doughty, 2019; Skehan, 1998, 2002). Skehan (1998, 2002) also noted that working memory is implicated in all stages of second language acquisition (language input, central processing, and language output) and in various language operations mechanisms (input processing, noticing, pattern recognition, pattern restructuring and manipulations, and pattern control). In a similar vein, the importance of working memory as a language aptitude variable is highlighted in the Hi-LAB (Doughty et al., 2007, 2019), where it is implicated in eight of ten measures: non-word span, anti-saccade, running memory span, task-switching, letter sets, paired associates, serial reaction time, and available long-term memory for synonyms.

Metacognition

Considered both as a trait and state attribute (Hong, 1998; Norman et al., 2019), metacognition was defined by Flavell (1979) as the “knowledge and cognition about cognitive phenomena” (p. 906). He further distinguished metacognition into metacognitive knowledge (knowledge of beliefs about cognitive aspects of people, tasks, or strategies), experiences (an individual’s cognitive or affective experiences accompanying a cognitive task), and strategies (methods an individual uses to control cognition as they carry out cognitive tasks).

Since the 1980s, there have been consistent research efforts to investigate the role of metacognition in L2 reading comprehension. Clearly related to Flavell’s (1979) ideas of metacognitive knowledge and experiences, some commonly investigated topics included L2 users’ self-perception as readers and L2 readers’ knowledge of text characteristics (e.g., Zhang, 2010). In addition, possibly due to its immediate relevance to reading assessment, much research attention has been paid to the knowledge and use of L2 reading strategies (e.g., Mokthari & Shorey, 2002; Schoonen et al., 1998).

While reading, metacognition allows the reader to determine the purpose of reading (e.g., skimming, scanning, and reading to learn), control reading rates accordingly (Carver, 1997), engage background knowledge to aid comprehension, monitor the quality of comprehension, and execute troubleshooting strategies when problems with comprehension arise. The earlier-introduced interactive-compensatory model of reading contends that such executive control processes play an important role among less proficient L2 readers because they deploy reading strategies and background knowledge to compensate for the lacking L2 knowledge and inefficient processing.
**Conative and Affective Variables**

**Motivation**

Motivation is a complex multidimensional construct and reading-specific motivation is a domain-specific construct distinct from motivation for other domains (e.g., math, science, or work; Wigfield, 1997). There is no single definition of reading motivation, but it includes such constructs as “goals, values, beliefs, and dispositions toward reading” (Grabe & Stoller, 2019, p. 83). In their seminal work on reading motivation, Wigfield and Guthrie (1997), postulated as many as 11 dimensions in L1 reading motivation and empirically supported the existence of all dimensions among Grade 4 and Grade 5 children. Wang and Guthrie (2004) revised this framework and proposed an eight-dimension model that covers intrinsic and extrinsic motivation. In a narrative synthesis of research, Schiefele et al. (2012) concluded that the distinction between intrinsic and extrinsic motivation plays a dominant role as a higher-order category, each of which subsumes several specific dimensions. The researchers further argued that genuine dimensions of reading motivation may include curiosity, involvement, competition, recognition, grades, compliance, and work avoidance.

Regarding the role of motivation in reading achievement, Schiefele et al. (2012) has found the following general trends: competence in reading comprehension is likely to be positively related to intrinsic motivation and negatively related to extrinsic motivation, and breadth or preference of reading may be more dependent on intrinsic than extrinsic motivation. However, a consensus has yet to be reached in many areas of inquiry including the relationship between reading motivation and reading amount, the ways in which reading motivation affects reading competence (e.g., directly, in a mediated manner, or bidirectionally), or possible factors that may mediate the relationship. Furthermore, how different dimensions of reading motivation relate to different reading outcomes needs to be examined (e.g., Wang & Guthrie, 2004; Wigfield & Guthrie, 1997).

Although gradually increasing, research on L2 reading motivation lags behind the large body of research on L1 reading motivation and on motivation for L2 learning/communication. With different empirical studies finding different components of L2 reading motivation in their L2 participants, a consensus has yet to emerge on the construct of L2 reading motivation. In the limited, extant research, the components of L2 reading motivation have often been related to intrinsic and extrinsic motivation (Komiyama, 2013). In identifying the construct of L2 reading motivation, one important question concerns how students’ L1 reading motivation relates to their L2 reading motivation (or whether the construct of reading motivation is common across L1 and L2). At present, available studies do not seem to support strong relationships between L1 and L2 reading motivation (Kim, 2010; Takase, 2007). This result, if generalizable, gives an impetus to conceptualizing L2 reading motivation as a language-specific construct.

**Anxiety**

Foreign language (FL) anxiety (Horwitz, 1986; Horwitz et al., 1986) and foreign language reading anxiety (FLRA) have been investigated in relation to L2 reading (Saito et al., 1999). FL anxiety is a type of situation-specific anxiety in FL classroom settings especially concerning oral language skills. FLRA is a more narrowly focused, skill-specific anxiety observed in the context of L2 reading. Saito et al. (1999) argued that two factors that are unique to reading may arouse FLRA: unfamiliar scripts/writing systems and unfamiliar cultural materials. Whether FLRA is a unique construct distinguishable from general FL anxiety is a critical question that needs to be empirically tested. Research has reported significant correlations between FLRA and general FL anxiety: $r = 0.64$ in Saito et al. (1999), and $r = 0.62$ in Zhao et al. (2013). However, the magnitude of correlations suggested that, although FLRA is related to FL anxiety, it is a distinct construct from FL anxiety.

Different views are expressed about the possible mechanism through which FL anxiety or FLRA negatively affects L2 reading performance. The original account is that anxiety may induce
IDs for L2 Reading

behaviors or mental states that could impede L2 reading (e.g., lack of concentration, forgetfulness, and the avoidance of learning opportunities) (Horwitz et al., 1986). From a more cognitive viewpoint, anxiety may restrict working memory capacity by reducing the attention that could otherwise be allocated to the reading task (Moran, 2016; Zhao et al., 2013). Further, Sparks and Patton (2013) argued that FL anxiety is associated with more general language skills such as L1 skills and L2 aptitude (e.g., low L1 skills inducing high FL anxiety) and that the negative relationship between FL anxiety and FL learning is the reflection of the relationship between language skills and FL learning.

Research

Evidence

Phonological Awareness

It is now well-established that phonological awareness is a strong predictor of word reading abilities and, consequently, that of reading comprehension in L1 (Scarborough, 2001; Stanovich, 1986). As such, the central question regarding phonological awareness in the context of L2 reading is not so much whether phonological awareness in L2 contributes to L2 reading as to 1) whether phonological awareness in L1 is related to L2 reading (cf. linguistic coding difference hypothesis), and if so, 2) in what mechanism it is. Related to the first question, Melby-Lervåg and Lervåg (2011) meta-analyzed the correlation between L1 and L2 phonological awareness. In this study, a total of 16 effect sizes drawn from studies involving children aged between four and 14 were meta-analyzed to yield a significant and large overall correlation: $r = 0.60, 95\% \text{ CI (0.48, 0.69)}$. This provided support for the linguistic coding differences hypothesis as well as the crosslinguistic effects between L1 and L2 phonological awareness.

Whether the strong relationship between L1 and L2 phonological awareness implies a strong contribution of L1 phonological knowledge to L2 reading is less clear. As part of their investigation on the crosslinguistic contribution of morphological awareness to L2 reading among 66 Arabic L1, English L2 high school students, Kahn–Horwitz and Saba (2018) found that L1 phonological awareness was a significant contributor to L2 word reading and reading comprehension. Similarly, Lindsey et al. (2003), in their longitudinal study tracking 249 Spanish L1, English L2 children from kindergarten to Grade 1, found a significant correlation between earlier L1 phonological awareness and later L2 reading comprehension. On the other hand, Gottardo and Mueller’s (2009) structural model of reading on Spanish L1, English L2 children showed that the direct path from L1 phonological awareness to L2 word reading was not significant. Only L2 phonological awareness had a direct path to L2 word reading, which then had a significant direct path to L2 reading comprehension. In sum, the results of the studies reviewed here indicate that while it is clear that the L1 and L2 phonological relationship is clearly strong, the extent to which L1 phonological awareness influences L2 reading, be it word- or passage-level, remains inconclusive.

Working Memory

Central questions regarding the role of working memory in L2 reading can be summarized as follows: 1) how strong is the relationship between working memory and L2 reading comprehension?; 2) do variables such as different dimensions of working memory or different measurement characteristics of working memory moderate the relationship between working memory and L2 reading comprehension?; and 3) how do features of the reading process (e.g., reading a hypertext, electronic reading tools that allow vocabulary search) interact with working memory?

Regarding the first and the second question, two meta-analyses by Jeon and Yamashita (to appear) and Shin (2020) provide relevant findings. Although differing in study features such as
The central questions on metacognition in the context of L2 reading vary considerably in their scope. The first type of questions is of a broader scope and examines issues such as the relationship between metacognition and L2 reading comprehension (e.g., Jeon & Yamashita, to appear) and the role of metacognition as a predictor of L2 reading comprehension (Guo, 2018; Schoonen et al., 1998; Van Gelderen et al., 2004). The second type of questions investigates the use of specific dimensions of metacognition such as reading strategies, self-assessment of reading, and comprehension monitoring (Barnett, 1988; Brantmeier & Vanderplank, 2008; Morrison, 2004; Shang, 2010).

Regarding the first question, research shows that metacognition overall has a significant and moderate relationship with L2 reading comprehension ($r = 0.330$, 95% CI [0.083−0.539]; Jeon & Yamashita, to appear). Further, Schoonen et al. (1998) and Van Gelderen et al. (2004) tested a model...
of L1 and L2 reading which included key linguistic knowledge variables and metacognition among Dutch secondary school students, and they reported that metacognition was a strong predictor of L2 reading comprehension, especially among older L2 readers (Grade 8 in Van Gelderen et al. [2004] and Grade 10 in Schoonen et al. [1998]). These results potentially indicated the increasingly important role of metacognition as a function of age. Additional support for this conjecture is found in Guo’s (2018) study where metacognitive knowledge indirectly predicted L2 reading through L1 reading ability and L2 language proficiency among university-level Chinese L1 students.

Research investigating the second question yielded many useful findings. Barnett (1988) investigated the relationship between reading strategies (actual and perceived reading strategies pertaining to using local and global contextual clues to help comprehension) and L2 reading outcomes among 278 university students studying French as a foreign language. Her results showed that students’ perceived and actual use of reading strategies were significantly correlated. In addition, both the perceived and actual use of strategies were significantly and moderately correlated with L2 reading comprehension. Shang (2010) investigated which types of reading strategies were most commonly used among 53 Taiwanese university EFL (English as a Foreign Language) students. The results showed that metacognitive, compensation, and cognitive strategies were the most frequently used. In addition, the use of reading strategies was also significantly correlated with participants’ self-efficacy as readers although it did not significantly correlate with actual L2 reading comprehension.

Brantmeier and Vanderplank (2008) and Morrison (2004) examined readers’ self-assessment of reading comprehension and how it relates to actual reading comprehension in L2. The first study, which included 359 university students who were advanced students of Spanish as a foreign language, investigated whether the self-assessment of L2 reading using a self-report instrument can serve as a useful placement test. Study participants completed a self-assessment survey about their reading abilities before and after completing an L2 reading comprehension test featuring various item types (e.g., recall, sentence completion, and multiple choice). Results showed that the self-assessment both before and after the reading comprehension test was positively associated with actual reading comprehension. In sum, the results indicated that advanced L2 readers are capable of correctly assessing their actual reading performance. Morrison (2004) investigated the relationship between comprehension monitoring in the form of error detection (e.g., detecting a graphemic error in a word and contradictions in the passage), the self-assessment of comprehension, and L2 reading proficiency among university students studying French in Canada. The results showed that error detection ability and L2 reading were significantly and strongly correlated. Both variables were also positively correlated with the self-assessment of reading comprehension.

Motivation

Key questions about motivation in L2 reading research have centered on the relationship between L2 reading motivation and various L2 reading outcomes. Takase (2007) reported that intrinsic motivation in L1 and L2 reading explained the amount of L2-English reading among L1 Japanese students. Kim (2010) found that scores on three dimensions of L2 reading motivation (i.e., learning goal-oriented motivation, intrinsic motivation, and avoidance) correlated with L2 (English) proficiency among L1 Korean university students. Lin et al. (2011) examined the predictability of different dimensions of reading motivation on L1 and L2 reading comprehension among Grade 5 L1 Chinese and L2 English children. Results showed that recreation, a type of intrinsic motivation, was a significant predictor of L1 reading, whereas instrumentalism, a type of extrinsic motivation, was a significant predictor of L2 reading. Dhanapala (2008) found different patterns of correlations between reading motivation and reading comprehension in L2 English between Sri Lankan and Japanese university students. Broadly, reading comprehension was strongly associated with both intrinsic and extrinsic motivation in the Sri Lankan group, whereas the correlation was much weaker in both categories of motivation (especially extrinsic motivation) in the Japanese
group. Using a more rigorous statistical method, Dhanapara and Hirakata (2016) expanded on the Sri Lankan study and showed that L2 reading comprehension was predicted positively by intrinsic motivation and negatively by extrinsic motivation among Sri Lankan students. In sum, research into L2 reading motivation is still in its early stages with little consistency in results. Efforts should be continued and expanded to gain insights into this important construct for L2 reading development.

Anxiety

L2 researchers have been working on the construct of L2 reading-specific anxiety (FLRA) and the relationship between FL anxiety/FLRA and L2 reading performance. First, regarding the construct, Saito et al. (1999) originally conceptualized FLRA as a unidimensional construct based on the high internal consistency of their instrument with L1 English university students learning three FLs (French, Japanese, and Russian). However, subsequent studies that examined FLRA’s components statistically (e.g., factor analysis and principal component analysis) identified multiple components, with the exception of Hsiao (2002, cited in Hamada & Takaki, 2021) which found a single factor supporting the unidimensionality. Multiple components were found in different studies, such as familiarity with vocabulary and grammar, reading confidence and enjoyment, and language distance (Hamada & Takaki, 2021; Matsuda & Gobel, 2004); top-down (reader-related) reading anxiety, bottom-up (text-related) reading anxiety, and classroom reading anxiety (Zoghi & Alivandivafa, 2014); and worry about comprehension, lack of satisfaction in reading ability, unfamiliar phonics roles (Bensalem, 2020). Thus, there is no consensus on the underlying factors of FLRA. Different results are attributable to a wide range of variables including the instrument used to elicit FLRA, learners’ backgrounds (e.g., age, proficiency, L1-L2 pairing, or beliefs) and sociocultural and educational contexts.

The second line of research on the relationship between FLRA and L2 reading comprehension has also yielded a complex picture. Despite theoretical expectations, past studies reported mixed results with associations between anxiety and L2 achievement being negative, positive, or insignificant (Teimouri et al., 2019). Teimouri et al.’s (2019) meta-analysis synthesized 216 effect sizes from 105 independent samples (involving 19,933 participants) to examine the correlation between L2 anxiety and L2 achievement. This study included L2 anxiety of various types such as general FL anxiety, FLRA, and other skill-specific anxiety. The mean correlation of FLRA with reading was negative and moderate in size ($r = -0.38$). Potential moderating factors have been reported in individual studies. For instance, Mikami et al. (2018) found a significant negative correlation of FLRA when the test of L2 reading comprehension was challenging; the correlation was not significant when the test was not demanding. Hamada and Takaki (2021) reported different degrees of contribution to L2 reading proficiency among different components of FLRA: the strongest negative impact by lack of confidence in reading, followed by a perceived level of difficulty in reading, and anxiety in language distance between L1 and L2.

Data Elicitation

Overview: Study Features of L2 Reading Research

Historically, the majority of study participants in L2 reading research have been older L2 learners in a foreign or second language setting (Jeon & Yamashita, 2021). This trend, however, is changing with the rapid growth of bilingual literacy research involving younger L2 learners (e.g., Lam et al., 2012; Pasquarella et al., 2011). Studies involving variables that have stronger implications in the earlier stages of reading development, such as phonological awareness, also predominantly involve kindergarteners or children in early elementary grades (e.g., Chen et al., 2012). In addition, studies that longitudinally track the predictability of various reading components
on reading comprehension also often follow a group of younger readers from an early age (e.g., Grant et al., 2012).

In terms of research design, the studies are primarily correlational, and experimental studies are restricted to the investigation of the effectiveness of strategy instruction. In a typical correlational study, the researcher obtains two sets of scores: one for the independent or predictor variable(s), such as motivation, and one for reading comprehension. Then, correlational analyses such as regression analysis, path analysis, or structural equation modeling (SEM) are conducted to ascertain whether the ID variables are significantly predictive of reading comprehension. On the other hand, intervention studies investigate whether focused instruction on certain variables would enhance reading outcomes. To our knowledge, intervention studies targeting ID variables to improve L2 reading comprehension are rare with the exception of those investigating the effects of reading strategy training on L2 reading; these studies typically involve a period of explicit instruction on cognitive (i.e., guessing the meaning of an unknown word) and metacognitive (e.g., setting reading goals and monitoring comprehension) strategies that are believed to facilitate reading comprehension and assess the effectiveness of treatment by comparing reading comprehension between an experimental group and a control group. For interested readers, we suggest Taylor et al.'s (2006) meta-analysis on treatment effect studies on L2 reading strategy use.

**L2 Reading Comprehension**

A variety of tests with a wide range of item types have been used to assess L2 reading comprehension. Jeon and Yamashita’s (2021) review of L2 reading assessment reports that commonly used standardized tests of reading comprehension (e.g., TOEFL-PBT, Nelson–Denny reading test, Woodcock language proficiency battery–passage comprehension subtest) predominantly rely on expository reading passages, multiple-choice questions (MCQ), and cloze items. In contrast, researcher-made tests, while they also often include expository reading passages, utilize a wider variety of item types such as recall (oral and written, using L1 or L2), summary completion, open-ended questions, or sentence completion. As Jeon and Yamashita (2021) noted, however, the seemingly uniform use of fewer item types by standardized tests must not be automatically interpreted to mean uniformity in target constructs; Keenan, Betjemann, and Olson (2008) showed that the scores of four popular standardized reading tests, namely, the Gray oral reading test (GORT), the qualitative reading inventory (QRI), the Woodcock–Johnson passage comprehension (WJPC) subtest, and the reading comprehension test of the Peabody individual achievement test (PIAT) correlated only moderately, suggesting that these tests tap into different cognitive abilities. A closer examination of the test data indeed indicated that listening comprehension accounted for the most individual variance in the GORT and QRI data, while decoding accounted for the most variance in the WJPC and PIAT data. Given this result, it is needless to say that there would be an even larger variation in the constructs captured by researcher-made reading tests that are much diverse in item types and administration methods. Therefore, it is crucial for the consumer of reading tests to carefully examine the target construct of the test and assess the potential effects of test administration methods as they determine which test is to be used for their study. For more detailed information about test selection, we direct the interested readers to the “Recommendations for Practice” section of Jeon and Yamashita (2021).

**Phonological Awareness**

Both in researcher-developed tests (e.g., Abu–Rabia & Sanitsky, 2010; Lefrançois & Armand, 2003) and standardized tests (e.g., Chen et al., 2012; Swanson et al., 2011), phonological awareness is commonly assessed through tasks involving manipulations of sound structures, for example phoneme deletion, insertion, substitution, identification, categorization, comparison, permutation, and
Eun Hee Jeon and Junko Yamashita

syllable awareness. Jeon and Yamashita’s (to appear) review of L2 reading studies investigating phonological awareness of the past three decades showed that standardized tests and researcher-made tests are used in approximately equal measure and that the two most frequently used standardized tests of phonological awareness were Rosner’s (1993) test of auditory analysis skills and comprehensive test of phonological processing (CTOPP; Wagner et al., 2013).

Working Memory

The complex and multicomponential nature of working memory is well-reflected in the diversity of its measurement instruments. Jeon and Yamashita’s (to appear) review of working memory measures used in L2 reading research of the past three decades shows that working memory tests vary in the modality of input (e.g., visual, numerical, and verbal), testing language (e.g., L1 and L2), and target subcomponents (e.g., storage, processing, and executive control). Span tests, one of the most frequently used tests of working memory, also feature a range of options: reading span, digit span, color span, counting span, operation span, conceptual span, and listening span. Regarding the potential effects of methodological features of reading span tasks, Shin’s (2020) meta-analysis offers some useful pointers, in addition to the effect of testing language and that of recall order of final words (correct or random) as mentioned in an earlier section; working memory assessed through complex span measures that assess both storage and processing tend to be a better predictor of L2 reading comprehension.

Metacognition

At present, measurement of metacognition in L2 reading research relies heavily on self-report type instruments with a disproportionate interest in reading strategies: of the 11 individual studies published between 1988 and 2017 reviewed in Jeon and Yamashita (to appear), five of them included a self-report survey assessing the knowledge and/or use of reading strategies (Abu-Rabia & Sanitsky, 2010; Barnett, 1988; Guo & Roehrig, 2011; Shang, 2010; Shiotsu, 2010). A minority of studies have assessed the actual performance of metacognitive constructs such as comprehension monitoring or actual strategy use while reading (e.g., Barnett, 1988; Morrison, 2004), but such line of research has clearly been discontinued since the mid-2000s with no clear reason.

Motivation

Motivation is commonly assessed through self-report questionnaires with a Likert scale. Davis et al. (2018) reviewed this type of L1 reading motivation measures developed in and after 1990 and found as many as 16 measures (used either in the original form or with modifications). All instruments are available in English with a few also available in multiple languages; the majority of instruments are for elementary to secondary school students with only one for adult to senior readers. We do not see this level of variety in L2 studies, which is due partly to a smaller amount of research compared to L1 studies. Rather, L2 researchers modify L1 reading motivation questionnaire items according to their research contexts and purposes. Among the L1 instruments, L2 researchers have consulted the motivation for reading questionnaire (MRQ) most frequently (Wang & Guthrie, 2004; Wigfield & Guthrie, 1997), presumably because it is one of the most comprehensive and influential measures in L1 studies (Komiyama, 2013).

Anxiety

A self-report questionnaire with a Likert scale is the most common way of measuring FL anxiety including FLRA. Teimouri et al., (2019) found as many as 25 different questionnaires in the sample of their meta-analysis. As a measure of FLRA, Saito et al. (1999) developed the foreign language reading anxiety scale (FLRAS). This questionnaire was modeled after the measure of general FL
IDs for L2 Reading

anxiety, the foreign language classroom anxiety scale (FLCAS; Horwitz et al., 1986). FLRAS contains 20 items on a five-point Likert scale running from “strongly agree” to “strongly disagree” and elicits students’ self-report of anxiety over various aspects of reading and their perceived difficulty of reading compared to other language skills. High internal consistencies of FLRAS, as well as FLCAS, are reported (Teimouri et al., 2019). Although reading researchers tend to use these time-tested questionnaires (with some modifications where necessary), new self-report instruments have also been produced in efforts to create an instrument that better suits each specific context, e.g., the English as a foreign language reading anxiety inventory (EFLRAI; Zoghi & Alivandivafa, 2014) or the second language reading anxiety scale (SLRAS; Cheng, 2017). These new instruments tend to adopt the same format as FLCAS/FLRAS, which could be viewed positively in light of the reduction of possible effects of measurement format when research results on FLRA are compared across different studies. The new instruments were validated by their original authors, but validation efforts should be continued by different researchers in different contexts to see the applicability of the new instruments outside the original venues.

Practical Applications

The importance of phonological awareness, especially in the early stages of reading development, indicates that L2 instruction should start with ample oral language activities to help learners internalize basic units of L2 spoken language (e.g., phonemes, onset-rimes, and syllables). Although L1 and L2 phonological awareness are associated, when L1 and L2 are different in fundamental phonological units, learners need to acquire L2-specific phonological awareness. Phonological awareness and reading ability develop reciprocally, i.e., phonological awareness not only supports reading but also further develops through reading. Therefore, teachers should include oral reading activities to help learners acquire L2 grapheme–phoneme correspondences and fine-tune or improve their initial phonological awareness by reading.

Working memory is often regarded as a domain-general (then, language-free) system. However, as reviewed above, research suggests the mediating effect of L2 proficiency on the relationship between working memory and L2 reading (Shin, 2020). This finding has implications for helping learners with lower working memory capacity. Since the effect of working memory training is not yet supported (Melby-Lervåg & Hulme, 2013), continued efforts to improve L2 language abilities seem one of the surest ways to compensate for a lower working memory capacity. With stronger L2 proficiency (knowledge and fluency in L2 processing), learners can store, manipulate, and process L2 elements more efficiently without unduly sacrificing their working memory capacity. Emerging research on the impact of working memory on digital text reading calls for teachers to pay attention to learners’ problems when they are engaged in various digital/multimedia text readings. The difficulties may be unique to the newer reading modes apart from traditional paper-text reading. Although more research is needed, teachers could assist learners by, for example, suggesting efficient strategies to cope with newer reading modes.

Finally, we would like to address the practical implications of other ID variables that we discussed in previous sections. Regarding metacognition, training metacognitive strategies (e.g., monitoring comprehension, identifying reading difficulties, planning and forming goals for reading, and reading selectively according to the goals) have long been recommended in reading strategy training. Grabe and Stoller (2019) distinguished teaching individual reading strategies from training strategic readers. Broadly, the former, which is more conventional, is more like knowledge building (introducing strategies to students, often one strategy at a time in decontextualized contexts), whereas the latter refers to the actual training of strategies (activities to use multiple strategies in actual reading for comprehension and to recycle strategies in different combinations in different reading contexts). The authors strongly recommend using the latter method for strategy training. Due partly to a small amount of research, we do not yet know what kinds of L2 reading motiva-
tion facilitate L2 reading achievement. The relationship between motivation and L2 reading is likely to be affected by contexts. Practitioners should pay close attention to individual students in their specific contexts and encourage them to achieve their best performance. Although consensus has not been achieved on what causes anxiety in L2 reading, research has supported the negative relationship between anxiety and L2 reading achievement. Therefore, teachers should make efforts to reduce learners’ anxiety in reading contexts.

Future Directions

We would like to identify the following areas for further research. Phonological awareness has been one of the most extensively researched areas with largely established measurement methods. A unique question in L2 reading is, as discussed earlier, whether and how L1 phonological awareness affects L2 reading. This relationship is likely to be mediated by various factors including, but not limited to, L1–L2 differences in phonological structures and orthographic depth, age, instruction, and L2 exposure. Considering the importance of phonological awareness for reading, future studies need to focus on the mediating factors to understand how and when learners could benefit from their L1. Although there are many areas of investigation, the effects of working memory on digital/electronic reading create an important question for future reading research. Since a wide variety exists both in the type of new reading modes and in the measurement of working memory, efforts to conduct individual studies are necessary before we can synthesize research findings and obtain more generalizable results. Metacognition is a broad concept that eventually enables self-regulation in reading processes. The research method discussed above indicates that the current research practice limits its measurement to explicit knowledge of metacognition, especially reading strategies. Whether readers put their knowledge to actual use, and whether and how the actual use of metacognitive regulations relate to L2 reading achievement, have not been sufficiently understood.

Finally, despite the frequent use of self-report questionnaires, L2 research has not yet established a common instrument to measure L2 reading motivation and anxiety. This tendency is especially strong in the motivation field. Different questionnaires are likely to tap into different underlying constructs, and if different studies measure different constructs, synthesizing research findings becomes difficult. Having said that, however, it is likely that motivation and anxiety have context-specific dimensions, which is why researchers continue creating new instruments in their venue. One way to solve the dilemma between generalizability (universality) and context-specificity may be to have two layers in the designing feature of questionnaire instruments. The higher level consists of universal categories such as intrinsic versus extrinsic motivation, and the lower level encompasses various context-specific dimensions in each higher-level category. This design feature may enable researchers to measure context-specific dimensions, summarize them into higher-level categories to examine the universal applicability of their findings, and finally contribute to more general theory building.

References


IDs for L2 Reading


