

*Article*

## **Cognitive Processes and Foreign Language Reading: Investigating Students' Test Taking Behaviour with a Reflective Questionnaire**

**Klaus Siller**

Salzburg College of Education, Austria

**Andrea Kulmhofer-Bommer\***

Federal Institute of Quality Assurance of the Austrian School System, Austria

**Ulrike Kipman**

Salzburg College of Education, Austria

Received: 15 September 2021/Accepted: 31 January 2022/Published: 30 March 2022

### **Abstract**

Commonly, models of reading emphasize the relevance of cognitive processes and metacognitive strategies when it comes to reading comprehension. Although these models are derived from research into first language (L1) reading, they often serve as the basis for operationalizing foreign language (FL) reading comprehension tests. This is also true for the FL Reading Test (E8 Reading Test) administered in 2013 and 2019 in Austria. However, little is known about test takers' behaviour regarding their cognitive processes and metacognitive strategies when taking the test. Undoubtedly, such information could be useful for the FL classroom and to inform foreign language reading instruction as well as test development. Therefore, this paper investigates which cognitive processes and metacognitive strategies students (n=106) apply when taking the E8 Reading Test. Data was collected using a reflective questionnaire. The results show that compared to weaker readers, strong readers apply the expected cognitive processes more frequently. There is no statistically significant difference, however, between stronger and weaker readers regarding metacognitive strategies applied. Furthermore, the data revealed that some good readers arrive at the correct answer by making use of different/other cognitive processes and metacognitive strategies than expected. These findings emphasize the importance of explicit, guided foreign language reading instruction focussing not only on the product of reading (comprehension) but also on the processes involved. However, more research is needed to better understand what the absence or presence of skills and strategies mean regarding individual learner abilities.

### **Keywords**

Foreign language reading, foreign language instruction, assessment, metacognitive strategies, cognitive processes

---

\*Corresponding author. Email: [andrea.kulmhofer-bommer@gmx.at](mailto:andrea.kulmhofer-bommer@gmx.at)

## 1 Introduction

Modern foreign language (FL) instruction aims at putting equal emphasis on all four skills – reading, writing, listening, and speaking – by acknowledging that the four skills do not only support each other but reflect real language use. Thus, modern FL instruction aims to expose FL learners to authentic language and emphasise meaning (Garcia, 2021). This approach, however, assumes that reading skills are successfully transferred from the learners first language (L1) onto the FL and thus not enough emphasis is given to (basic) reading skills and strategies in FL instruction (Villacañas de Castro, 2016).

There are nonetheless theories which state that the positive transfer of reading skills can only occur if the reader has already developed a certain level of experience with reading in the L1 (Geva et al., 2019) and a certain competence in the foreign language (Linguistic Threshold Hypothesis: Cummins, 1979). Another factor impacting on how easy or challenging reading in a FL is lies in the orthographic depth of the language one learns to read in (Orthographic Depth Hypothesis: Katz & Frost, 1992) making English a particularly challenging language to be decoded for students whose L1 lies on the other end of the transparency/opacity continuum such as German, Italian or Finnish (Calero & Calero-Pérez, 2021). Taking into consideration that large scale international assessments such as PISA (OECD, 2019) tend to reveal that a significant number of students aged 14 do not have sufficient literacy skills to successfully participate in academia or on the job market (OECD, 2019), the importance for reading instruction in the higher grades (grades 5 to 8) becomes even more evident. Low reading skills in L1 negatively impact FL reading development as learners cannot rely on strategies and skills to master FL reading (Alderson, 1984; Kahn-Horwitz & Saba, 2018; Koda, 2005). This makes reading instruction a necessity in the FL classroom if teachers want to provide equal opportunities to all students (Verhoeven, Perfetti & Pugh, 2019; Villacañas de Castro, 2016). Furthermore, decades of research have shown that explicit reading instruction in the FL classroom does not only impact FL learning but also positively impacts L1 reading because students develop metacognitive strategies and cognitive processes that can then be transferred to L1 reading (Kormos, 2020; Sparks, Schneider & Ganschow, 2002).

Traditionally, FL learners' reading skills are assessed with FL reading assessments, such as reading comprehension tests. These tests may discriminate well between stronger and weaker readers, information that is usually based on quantitative test results. Whether such test results provide enough information for teachers in the classroom, however, needs to be determined. To improve classroom procedures and learning techniques, teachers need rather more specific information: for example, by gaining qualitative insights into test takers actual processing when answering test items. Against this backdrop, the current paper investigates what Austrian 8<sup>th</sup> grade students (age 14) actually do when completing FL English reading comprehension tasks. Conclusions are drawn about the impact of these findings on FL English reading instruction in Austria.

## 2 Literature Review

### 2.1 Models of (foreign language) reading

Reading is a very complex process and an enormous amount of research into reading comprehension is available. Despite this, no definitive theory of reading has emerged thus far and probably never will. For example, terms such as reading strategies and reading (sub)skills have been discussed as being the basis for FL reading comprehension in the literature (see, for example, Grabe, 2000; Williams & Moran, 1989). Quite often these terms have been used interchangeably. Williams and Moran (1989) point out, however, that reading strategies are applied consciously by the reader to comprehend printed text. These comprise cognitive strategies, supporting the language learning process (Singhal, 2001), and metacognitive strategies, supporting the comprehension process by planning for organization, selective attention, goal setting, monitoring and self-assessment (Ali & Razali, 2019; Muhid et al. 2020). Reading

(sub)skills, on the other hand, are automatized and applied unconsciously, for example decoding skills, linguistic knowledge, vocabulary background knowledge, metacognitive knowledge, discourse structure knowledge and skills monitoring (Grabe, 2000; Grabe & Stoller, 2011).

There have been many attempts to provide frameworks to classify reading ability in more detail, with various models of reading described in the literature. Traditionally, process models and componential models have been used to describe the reading process in more detail (Urquhart & Weir, 1998). These models comprise so-called bottom-up and top-down processes that are involved in reading and that describe “how words are recognised, how long they are kept in working memory, when syntactic processing begins, and so on” (Urquhart & Weir, 1998, p. 39). Bottom-up processes are text-driven, that is, the information necessary to comprehend a text is to be found in the printed words of the text only. To make meaning of the provided information, a reader has to decode letters, words, phrases, sentences and recognize grammatical and other linguistic cues (see, for example, Alderson, 2000). Top-down approaches, on the other hand, view the comprehension of written texts as reader-driven, that is, the reader includes knowledge of the world, experience with texts and other schemata into the comprehension process (Goodman, 1967; Smith, 1971).

Both bottom-up and top-down models have been criticized in the literature, as research has shown that reading requires both top-down and bottom-up strategies to be successful. Bottom-up models have been criticized for their view of reading as a linear and unidirectional process (Bernhardt, 1991; Rumelhart, 1977), top-down models for failing to explain how new meaning is assimilated with prior meaning (e.g., background knowledge). Grabe and Stoller (2011, p. 26), for example, criticize extreme interpretations of top-down models when they say that “there is a question about what a reader can learn from a text if the reader must first have expectations about all the information in the text”.

Currently, then, more interactive models of reading are preferred by researchers. This synthesis of existing views on bottom-up and top-down processes is represented, for instance, in Scarborough’s *Reading Rope Model* (Scarborough, 2001), Perfetti’s *Reading Systems Framework* (Perfetti & Stafura, 2014) or Khalifa and Weir’s (2009) *Cognitive Processing Model for Reading Comprehension*. Although describing the processes of reading for native speakers, these models are also widely used in FL or second language learning contexts. For instance, Khalifa and Weir (2009) argue that their model can be used for creating more authentic L2 reading tasks, a fact that has made their model very popular in the context of operationalizing reading tests. Khalifa and Weir (2009) distinguish bottom-up processes, including different types of reading and the cognitive processes these may invoke, and top-down processes, that is a reader’s knowledge base necessary to comprehend a written text. Depending on the type of reading used by the reader (expeditious and careful reading at a global or local level), various cognitive processes are triggered. These processes range from lower-order processes, such as decoding, to higher-order processes, for example creating both a text-level and intertextual representation of a single text or across texts, respectively. Additionally, knowledge that supports the comprehension of meaning is identified. It comprises, for example, lexical and syntactic knowledge, topic knowledge, general knowledge of the world and text structure knowledge. Depending on the reading skills of a reader, the knowledge base is used to either support decoding processes or enrich understanding of a text.

## 2.2 The E8 reading test

In Austria, it is only recently – that is in 2009 – that national educational standards for different subjects and different stages of schooling have been implemented. These standards provide the legal basis for an external assessment of Austrian students’ competences. The current paper focuses on the national reading test for English as a first foreign language for lower secondary level after eight years of schooling (E8 Reading Test). (For a detailed description of the development of the E8 Reading Test see Siller &

Kulmhofer, 2018). In the context of the E8 Reading Test, it became evident that cognitive processes and metacognitive reading strategies, as defined in Khalifa and Weir's (2009) reading model, had a significant impact on the difficulty of the items used in the test (Siller 2020; Siller & Kipman, 2018). Item difficulty was measured by calculating the proportion of test takers answering an item correctly; the higher the number of test takers who answered an item correctly, the easier the item.

In summer 2019, a reading model for the operationalization of prospective E8 Reading Tests was developed by the authors of this article. The model was reviewed by four subject experts (one teacher, two teacher trainers in EFL methodology and one language testing expert) and then used to analyze 48 items. The items were taken from the item pool of the Federal Institute for Quality Assurance of the Austrian School System (IQS; Institut für Qualitätssicherung im österreichischen Schulwesen), the institute responsible for the design and administration of the nation-wide standard tests (for more details see Siller, Kulmhofer-Bommer & George, 2020). The review process confirmed that, on the one hand, items in the E8 Reading Test are supposed to elicit metacognitive reading strategies. Test takers should apply expeditious local reading strategies, for example when they scan/search for explicitly mentioned information within a clause or sentence. Some items in the test will ask the readers to identify explicitly mentioned information within a single clause or across adjoining sentences, thus using a careful reading strategy at the local level. More complex strategies comprise expeditious global reading, with readers having to skim texts to identify implicit information scattered across adjoining sentences or the whole text, and careful global reading, asking for the identification of explicit and/or implicit information scattered across the whole text.

On the other hand, items in the E8 Reading Test are supposed to elicit cognitive processes of different complexity. For example, a rather simple process of decoding is applied when test takers have to identify identical words/phrases in the stimulus text, question(s) and/or answer option(s). A more complex process, establishing propositional meaning, is used when readers identify synonyms and/or paraphrases in the stimulus text, question(s) and answer option(s). Even more complex processes include inferencing, for example when inferring meaning at the word-, phrase- and/or sentence-level, and building a mental model, when readers have to differentiate between the importance of ideas and relate ideas from the text to each other. Furthermore, both metacognitive strategies and cognitive processes are influenced by an individual test taker's knowledge, in the E8 context most presumably their syntactic, lexical and topic knowledge.

Generally speaking, the E8 Reading Test discriminates well between stronger and weaker readers. However, in order to improve classroom procedures and learning techniques, teachers need more specific information, for example, on how individual test takers actually approach and answer reading items. In this context, the current paper investigates what cognitive processes, metacognitive strategies and what knowledge test takers actually apply when answering test items in the E8 context.

The overarching research question in the current study is therefore: **How do successful and unsuccessful readers differ in their approaches when answering reading items?** Specifically, the study is pursuing the following six hypotheses:

1. Do self-reported good, average, and weak readers differ in their ability to locate the necessary information, their use of metacognitive reading strategies, their application of cognitive processes and their use of knowledge?
2. Are there any gender-specific differences when solving reading tasks?
3. Do successful test takers locate the same text passage(s) that content experts think contain(s) the information necessary to answer a reading test item?
4. Do successful test takers apply the metacognitive reading strategies content experts think an item elicits?
5. Do successful test takers apply the cognitive processes content experts think an item elicits?
6. Do successful test takers use the knowledge that content experts think is useful when solving a reading test item?

## 3 Method

### 3.1 Participants

The test takers participating in this pilot study (N=106) were 8<sup>th</sup> grade students in an Austrian middle school. 57 participants were male (54%), 49 participants female (46%). 8<sup>th</sup> grade middle school students in Austria are about 14 years old. Information about students' age was not collected as the main criterion was for them to attend 8<sup>th</sup> grade and to receive instruction based on the 8<sup>th</sup> grade FL curriculum, the target group of the E8 Reading Test. All participants were enrolled in the same middle school in the north east of Styria but allocated to different strands (e.g., where the focus was on team sports, languages, arts) of the same high school. Information about the strand visited was not collected for the purpose of this pilot study as all middle school students are taught according to the same curriculum.

### 3.2 Research instruments

#### 3.2.1 Test tasks

In order to use valid test tasks, the research team gained access to the item pool of the Federal Institute for Quality Assurance of the Austrian School System (Institut für Qualitätssicherung im österreichischen Schulwesen; IQS), the institute responsible for the nation-wide standard tests. The tasks in the item pool of IQS are all linked to the E8 reading model and their quality has been approved in previous field studies. The tasks were selected on the basis that they:

1. included various text types used in the E8 Reading Test
2. dealt with rather familiar topics as identified by the Austrian curriculum for foreign languages (Rechtsinformationssystem des Bundes, 2012)
3. included various item formats used in the E8 Reading Test
4. included items that are supposed to elicit the metacognitive strategies outlined in the E8 reading model
5. included items that are supposed to elicit the cognitive processes outlined in the E8 reading model
6. included items that are supposed to elicit the necessary type of knowledge outlined in the E8 reading model

Finally, six tasks were selected. Due to the level of test takers in the E8 context (A2/B1 in terms of the *Common European Framework of Reference: Council of Europe, 2001*), all texts are written in rather simple language, have a very transparent structure and may include visuals to support comprehension. Each task used a slightly different item format; 24 items had to be answered:

*Task 1:* An email between two friends, accompanied by a multiple-choice item (three options, 1 correct answer). Test takers have to identify the correct purpose why this email was written.

*Task 2:* A magazine interview with a wildlife vet, accompanied by five items (extended matching headings). Five short paragraphs are provided and have to be matched to six interview questions (one question will not be needed). Test takers have to identify the main idea of each paragraph.

*Task 3:* A flyer advertising an outdoor week for 8 to 12-year-olds, accompanied by five open questions. Test takers have to identify some specific details from the flyer. (Grammar or spelling mistakes in the test takers' answers were not counted as incorrect, as long as the meaning was clear to the coders.)

*Task 4:* A simple recipe for a chocolate cake, accompanied by a multiple-choice item (three visuals as options, one correct answer). Test takers have to identify details in the text.

*Task 5:* Three short, illustrated magazine texts providing information about camping sites,

accompanied by four multiple-choice questions (four statements, three options per statement; more than one option may be correct). Test takers have to identify details in the texts.

*Task 6:* A literary text (an extract from *The Adventures of Tom Sawyer*), accompanied by an item set, that is, three multiple-choice questions (three options, one correct answer) and five true/false questions. Test takers have to identify the gist and some details in the text.

### 3.2.2 Reflective questionnaire

To investigate reading processes in reading tests, the literature (see, for example, Cohen, 1984; Gass & Mackey, 2000; Green, 1998) suggests experimental procedures such as stimulated recall or (verbal) protocol analysis where test takers verbally describe the linguistic processes they are engaged in when taking a test. However, the act of reporting may have a negative effect on the process of reading, as readers have to simultaneously verbalise their thoughts and solve test tasks. Especially, younger readers may lack both experience and language necessary for using this method. Therefore, for this pilot study a reflective questionnaire was designed that attempts to support test takers with self-reporting employed reading strategies, cognitive processes and the knowledge used when responding to the reading comprehension tasks. A reflective questionnaire aims at making thought processes visible that might otherwise be invisible to teachers and learners alike. According to Finlay (2008), learning involves thinking and doing. However, learners rarely think about what they are doing. Reflective questionnaires can support these thought processes and have to date gained attention in various fields including education, medicine and social work (eg., Nguyen et al., 2014; Tummons, 2011). The reflective questionnaire developed for the study at hand was discussed within the research team and trialled with a small number of learners from the target group (n=5). Based on the feedback received, some revisions were made, for example simplifying the language used in the questionnaire before it was used with a larger number of participants. (For a detailed description of the questionnaire and how it was used, see the section on data collection and coding below.)

## 3.3 Instrument development

First, a panel of judges was invited to analyse and discuss the selected tasks. The judges were two experts in language testing and teaching English as a FL and eleven master students at the College of Education in Salzburg. The master students were majors in teaching English as a foreign language and already had classroom experience as well as some basic experience in language testing. After a thorough familiarisation with the E8 reading model (see above) and two training rounds with sample tasks, the judges had to describe individually and for every task

1. where they found the information necessary to solve the task
2. which metacognitive reading strategy they think each item elicits
3. which cognitive process they think each item elicits
4. which type of knowledge they think would be helpful to answer each item

This procedure essentially follows the guidelines for developing a Q-matrix as described by, for example, Buck et al. (1998) or Li and Suen (2013). It is based on the assumption that judges familiar with the field know the best way to respond to an item. Collected responses are, however, highly subjective and depend on participants' experience with and knowledge about the reading test and its underlying construct. Thus, individual answers were discussed in a plenary session and a second round with the above questions was conducted. This increased agreement between judges, and it was possible to select the answers with the highest agreement between judges as references for further analyses.

Additionally, a cut-score had to be found to distinguish “successful” from “unsuccessful readers” in the target group. As mentioned above, the six reading tasks and 24 items had been trialled in a field study by the IQS. Each item was answered by a range of 194 to 766 students; therefore, data with regard to item facility, that is, the proportion of test takers selecting the correct answer to an item in the field test, was available for each item. Consequently, the IQS determined items with a facility index between 0.67 and 0.9 as easy, items with a facility index between 0.35 and 0.66 as moderately difficult, and items with a facility index lower than 0.34 as difficult. In a discussion amongst the judges, the decision was taken that “successful readers” would have to be able to answer more than 80% of the 24 items used.

Finally, a booklet was created for data collection within the target group. This booklet included a short demographic questionnaire about the participants’ personal background (e.g., age, grade, self-reported reading ability in EFL) and the six reading tasks with the reflective questionnaire after each task. Students had one regular class (50 minutes) to complete the demographic questionnaire, solve the reading tasks and fill in the reflective questionnaires. This process ensured that participants were not overburdened. The reflective questionnaire was structured and coded as follows:

*Location of necessary information:* Test takers had to identify and highlight the part of the text that included the information they used to answer the question. This way information was gained whether test takers and experts agree on where and what exactly the information is that enables test takers to answer an item correctly. Text passages were coded as (1) if they matched the passages the experts had selected, (2) if they did not match at all, (3) if nothing was highlighted (missing), (4) if the highlighted passage could not be matched to a corresponding item, and (5) if there was at least a partly match with the experts’ opinion.

*Information about how test takers approached the task:* This section sought information on whether test takers (1) first read the question and then the text, (2) first read the text and then the question or (3) did not answer the question at all.

*Metacognitive reading strategies:* This section sought information on whether test takers used global/local expeditious or global/local careful reading strategies to respond to the questions asked. Test takers were provided with five simple descriptions in their mother tongue, German, of how to approach a text and had to decide on the one that came closest to what they thought they did when solving the reading task. Their answers were coded as (1) if they did not read the full text but only a few words or phrases (expeditious local), (2) if they skimmed the full text (expeditious global), (3) if they skimmed the text to identify the place with the information they needed (both expeditious local and global), (4) if they only read the part with the necessary information slowly and very carefully (careful local), (5) if they read the full text slowly and very carefully, from the beginning to the end (careful global). Furthermore, code (0) was given when there was no answer and code (6) was given if more than one description was selected.

*Cognitive processes:* This section sought information on what test takers did to find the correct answer to each item. Test takers were provided with six simple descriptions in their mother tongue, German, of how to approach a text and had to decide on one. Their answers were coded as (1) if they searched for exactly the same words in the text that were used in the question (decoding), (2) if they searched for words or phrases in the text with a similar meaning to the words that were used in the question (establishing propositional meaning), (3) if they had to interpret/ “read between the lines” to find the correct answer (inferencing), (4) if they had to connect information from different parts of the text (building a mental model), (5) if they guessed the answer, and (6) if they said they could not answer the question. Furthermore, code (0) was given when there was no answer and code (7) was given if more than one description was selected.

*Test taker knowledge:* This section sought information on what type of knowledge test takers used to answer the questions. Test takers were provided with three simple descriptions in

their mother tongue, German, of what knowledge might be useful and had to decide on one. Their answers were coded as (1) if they used topical knowledge, (2) if they used grammatical knowledge, and (3) if they used lexical knowledge. Furthermore, code (0) was given when there was no answer and code (4) was given if more than one description was selected.

## 4 Data Collection

Data was collected in January and February 2020 by one teacher who had agreed to be the contact person at the school. She administered the data collection process according to an administrative protocol providing information about the overall procedure and details about students' tasks. The contact person was not an English teacher, ensuring that she could not help students in the process of completing the tasks and filling in the questionnaire. The data was collected in a total of four classes in four different data collection sessions. Verbal consent for the data collection was granted by the headmaster and the participants themselves. Furthermore, participation was voluntary, and participants could decline participation. To our knowledge, no participant refused. Thus, all students of one year group participated in this pilot study. Originally, it was planned to collect data at various schools. However, due to the COVID-19 pandemic and schools being in lockdown from March 2020, no further data collection was possible. For all statistical analyses, SPSS version 24 (2016) was used.

## 5 Results

### 5.1 Self-reported reading ability

As part of the demographic questionnaire, participants were asked to self-evaluate their reading abilities in English. They were able to choose between three options, coded as (1) excellent, (2) ok, and (3) not so good. These options were depicted as smiley faces. The three groups ("excellent" (N = 40), "ok" (N = 61) and "not so good" (N = 5) were compared based on their ability to locate the necessary information, their metacognitive reading strategies, their application of cognitive processes and their use of knowledge. The one-way MANOVA showed no statistically significant difference between the three groups,  $F(8,202)=.638, p=.745$ .

### 5.2 Gender differences

To check for possible gender differences, independent t-tests were conducted. An analysis of the self-evaluation of reading ability shows significant differences between the 57 male and 49 female students in their self-evaluation, which means that female students assessed themselves more positively than the male students. Statistics (mean, standard deviation (SD), and the t-value and p-value to show significant difference) are presented in Table 1. In a further analysis, female test takers were found to do significantly better than male test takers ( $p < .05$ ) when solving the reading tasks.

Table 1

*Gender Differences in Self-assessed Reading Ability*

	Male		Female		Difference	
	Mean	SD	Mean	SD	t	p
Self-assessed reading ability	1.54	0.54	1.82	0.57	2.54	.013*

Note: \* $p < .05$



### 5.3 Successful and unsuccessful readers

Based on the number of their correct answers, 106 test takers were separated into 64 successful readers (82% to 100% of answers correct) and 42 unsuccessful readers (17% to 79% of answers correct). These 64 successful and 42 unsuccessful readers were compared in t-tests regarding their ability to locate the necessary information, their application of metacognitive reading strategies and cognitive processes, and their use of knowledge. The results, presented below, show that there is a significant difference between successful and unsuccessful readers in their ability to locate the necessary information, their application of cognitive processes and their use of knowledge. No significant difference between successful and unsuccessful readers was found with regard to their application of metacognitive strategies. Statistics (mean, standard deviation (SD), the percentage (%) of successfully solved tasks, and the t-value and p-value to show significant difference) are presented in Table 2.

Table 2

#### *Differences Between Successful and Unsuccessful Readers*

	Successful readers			Unsuccessful readers			Difference	
	Mean	SD	%	Mean	SD	%	t	p
Location of information	.26	.29	26	.13	.17	13	2.82	.006**
Metacognitive strategies	.20	.201	20	.19	.199	19	.33	.744
Cognitive processes	.50	.27	50	.37	.27	37	.232	.022*
Use of knowledge	.71	.35	71	.53	.42	53	2.21	.030*

Note: \*p < .05, \*\* p < .01

Multiple Pearson correlations were conducted to assess the relationship between the variables. The analyses showed that the application of metacognitive reading strategies and the use of knowledge correlate positively with the ability to locate the necessary information. This means that the better students could apply metacognitive strategies and use their knowledge, the better they could locate the necessary information in the text. Furthermore, both cognitive processes and the use of knowledge correlate positively with metacognitive reading strategies. No significant correlations were found between the other variables. These variables were therefore largely independent of each other. The correlation coefficient (r) and the significance value (p) are presented in Table 3.

Table 3

#### *Correlations Between Variables*

	1		2		3		4	
	r	p	r	p	r	p	r	p
1. Location of information	-	-	.266**	.004	.095	.307	.257**	.005
2. Metacognitive strategies	.266**	.004	-	-	.181*	.050	.189*	.041
3. Cognitive processes	.095	.307	.181*	.050	-	-	.159	.086
4. Use of knowledge	.257**	.005	.189*	.041	.159	.086	-	-

Note: \*correlation significant at the .05 level (2-tailed); \*\*correlation significant at the .01 level (2-tailed)

## 6 Discussion and Implications

To answer the research questions posed, 106 8<sup>th</sup> grade students in an Austrian middle school were asked to complete a demographic questionnaire, to answer 24 reading items that were based on six different texts, and to complete a reflective questionnaire after each of the six reading tasks. Investigating the reading ability of the 106 learners with valid test tasks shows that readers who agree with content experts on:

1. where the information necessary to answer an item is located in a text
2. what cognitive processes should be applied
3. what knowledge is useful when answering a reading test item in the E8 context

are significantly more successful than those readers who do not agree. There is no statistically significant difference, however, between successful and unsuccessful readers regarding the metacognitive strategies applied. Also, there is no significant relationship between self-reported reading abilities and students' agreement with content experts regarding location of information, cognitive process, metacognitive strategy, and knowledge applied. This finding somewhat contradicts other research stating that how students judge their own abilities impacts their performances (e.g., [Asakereh & Yousofi, 2018](#)). Possible reasons for the contradictory findings could be the learners' age and language level. Further research is needed in this area.

For the FL teacher, the results indicate that there is a need to provide explicit, guided foreign language reading instruction in Austrian classrooms. Ticking reading comprehension questions and comparing solutions, a traditional and still very common approach in Austrian foreign language classes, does not meet the aims of active reading instruction. This does not mean that teachers have to actively teach reading theories and metalanguage for reading; but they should ask their learners to highlight any information in a text that they think helped them answer a specific comprehension question (locating necessary information). In addition, learners have to be shown different ways how to read a text (metacognitive reading strategies) and they should be able to explain what they did (cognitive processes) and what type of knowledge they used to answer a comprehension question (test taker knowledge). That way learners will be enabled to use metacognitive reading strategies, cognitive reading processes and their knowledge base more consciously and so become more successful readers.

An additional result of the pilot study showed significant gender differences with regard to both self-reported reading ability and actual reading performance. This result is in line with most studies focusing on gender differences in reading (see, for example, [Guiso et al., 2008](#); [Lynn & Mikk, 2009](#); [Reilly et al., 2019](#)). It should neither be argued here to apply different teaching styles to boys and girls nor to reintroduce same-sex schooling. However, as difficulties in reading have a detrimental effect on learning in general, it would be of interest to investigate the exact nature of such gender differences in further studies. Only then can teachers design targeted instruction to improve their learners' reading skills. This does not mean that gender differences should be highlighted at school, but, for example, offering a selection of texts or books might provide for individual interests better than making a class read the same texts and/or class readers all the time.

Furthermore, the data confirmed a result that, for example, [Sigott \(2004\)](#) described as *fluid construct phenomenon*. Some successful readers arrive at the correct answer by making use of different/other cognitive processes and metacognitive strategies than those expected by content experts. Also in this context further studies, for instance by using verbal protocols, would be of interest.

However, several limitations have to be taken into account when interpreting the results of this pilot study. First, due to the COVID-19 pandemic at the time of data collection it was not possible to gain data from various schools and more students. Thus, results should not be generalised to a student population different to the one described above. Second, the pilot study is of restricted scope, as only responses to 24

test items based on six texts were examined and the small sample size might have produced inconsistent results. Finally, expert judgements are highly subjective and there are doubts about the objectivity, validity and reliability of such expert ratings, as they very much depend on content experts' experience with and knowledge about the reading test and its underlying construct.

## 7 Conclusion

The main aim of this pilot study was to find out whether successful readers actually use the reading strategies and processes that are identified as supportive for reading comprehension in foreign language learning in the literature. 106 test takers had to respond to 24 reading test items and then complete a reflective questionnaire on where they found the correct answer (locating the necessary information), how they approached the texts (metacognitive strategies), and why they were able to answer the items (cognitive processes and useful knowledge). Results of the analyses suggest that those readers who apply the theoretical strategies, processes and knowledge that the items are supposed to elicit are more successful than those who do not apply them. It is hoped that the current research findings provide useful information for teachers to improve classroom procedures and learning techniques.

## Acknowledgement

The authors would like to thank Emily Kaiser for helping to structure and write the results section.

## References

- Alderson, J. C. (1984). Reading in a foreign language: A reading problem or a language problem? In J. C. Alderson & A. H. Urquhart (Eds.), *Reading in a foreign language* (pp. 1–24). Longman.
- Alderson, J. C. (2000). *Assessing reading. The Cambridge language assessment series*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511732935>
- Ali, A. M., & Razali, A. B. (2019). A review of studies on cognitive and metacognitive reading strategies in teaching reading comprehension for ESL/EFL learners. *English Language Teaching*, 12(6), 94–111. <https://doi.org/10.5539/elt.v12n6p94>
- Asakereh, A. & Yousofi, N. (2018). Reflective thinking, self-efficacy, self-esteem and academic achievement of Iranian EFL students in higher education: Is there a relationship? *International Journal of Educational Psychology*, 7(1), 68–89. <https://doi.org/10.17583/ijep.2018.2896>
- Bernhardt, E. B. (1991). *Reading development in a second language: Theoretical, empirical, and classroom perspectives*. Ablex.
- Buck, G., VanEssen, T., Tatsuoka, K., Kostin, I., Lutz, D., & Phelps, M. (1998). *Development, selection and validation of a set of cognitive and linguistic attributes for the SAT I Verbal: Analogy section* (Research Rep. No. RR-98-19). Educational Testing Service. <https://doi.org/10.1002/j.2333-8504.1998.tb01772.x>
- Calero, A., & Calero-Pérez, E. (2021). Orthographic depth and reading comprehension in Spanish. Opacidad ortográfica y aprendizaje de la comprensión lectora en español. *Ocnos*, 20(2), 33–42. [https://doi.org/10.18239/ocnos\\_2021.20.2.2274](https://doi.org/10.18239/ocnos_2021.20.2.2274)
- Cohen, A. D. (1984). On taking language tests: what the students report. *Language Testing*, 1(1), 70–81. <https://doi.org/10.1177/026553228400100106>
- Council of Europe (2001). *Common European framework of reference for languages: Learning*,

*teaching, assessment*. Press Syndicate of the University of Cambridge.

- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49(2), 222–251. <https://doi.org/10.3102/00346543049002222>
- Finlay, L. (2008). Reflecting on ‘Reflective practice’. Practice-based Professional Learning Paper 52, The Open University. Retrieved from <http://oro.open.ac.uk/68945/1/finlay-%282008%29-reflecting-on-reflective-practice-pbpl-paper-52.pdf>
- Garcia, M. (2021). Ten important aspects of a 21st century foreign language teaching approach. *European Journal of Foreign Language Teaching*, 5(5), 1-19. <https://doi.org/10.46827/ejfl.v5i5.3834>
- Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Lawrence Erlbaum Associates, Inc. <https://doi.org/10.4324/9781410606006>
- Geva, E., Xi, Y., Massey-Garrison, A., & Mak, J. Y. (2019). Assessing reading in second language learners: Development, validity, and educational considerations. In D. A. Kilpatrick, R. M. Joshi, & R. K. Wagner (Eds.), *Springer eBook Collection. Reading Development and Difficulties: Bridging the Gap Between Research and Practice* (1st ed., pp. 117–155). Springer. [https://doi.org/10.1007/978-3-030-26550-2\\_6](https://doi.org/10.1007/978-3-030-26550-2_6)
- Goodman, K. S. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist*, 6(4), 126–135. <https://doi.org/10.1080/19388076709556976>
- Grabe, W. (2000). Reassessing the term “interactive”. In P. L. Carrell, J. Devine, & D. E. Eskey (Eds.), *The Cambridge applied linguistics series. Interactive approaches to second language reading* (9th ed., Vol. 6, pp. 56–70). Cambridge University Press.
- Grabe, W., & Stoller, F. L. (2011). *Teaching and researching reading* (2nd ed.). *Applied linguistics in action*. Longman/Pearson. <https://doi.org/10.4324/9781315833743>
- Green, A. (1998). *Verbal protocol analysis in language testing research: A handbook. Studies in language testing: Vol. 5*. Cambridge University Press.
- Guiso, L., Monte, F., Sapienza, P., & Zingales, L. (2008). *Culture, gender, and math. Science*, 320(5880), 1164–1165. <https://doi.org/10.1126/science.1154094>
- Kahn-Horwitz, J., & Saba, M. (2018). Weak English foreign language readers: The cross-linguistic impact of morphological awareness. *Reading and Writing*, 31(8), 1843–1868. <https://doi.org/10.1007/s11145-017-9810-9>
- Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Advances in psychology: Vol. 94. Orthography, phonology, morphology, and meaning* (pp. 67–84). North-Holland. [https://doi.org/10.1016/S0166-4115\(08\)62789-2](https://doi.org/10.1016/S0166-4115(08)62789-2)
- Khalifa, H., & Weir, C. J. (2009). *Examining reading: Research and practice in assessing second language reading. Studies in language testing: Vol. 29*. Cambridge University Press.
- Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge applied linguistics. Cambridge University Press.
- Kormos, J. (2020). Specific learning difficulties in second language learning and teaching. *Language Teaching*, 53(2), 129–143. <https://doi.org/10.1017/S0261444819000442>
- Li, H., & Suen, H. K. (2013). Constructing and validating a Q-matrix for cognitive diagnostic analyses of a reading test. *Educational Assessment*, 18(1), 1–25. <https://doi.org/10.1080/10627197.2013.761522>
- Lynn, R., & Mikk, J. (2009). Sex differences in reading achievement. *Trames. Journal of the Humanities and Social Sciences*, 13(1), 3–13. <https://doi.org/10.3176/TR.2009.1.01>
- Muhid, A., Amalia, E. R., Hilaliyah, H., Budiana, N., & Wajdi, M. B. N. (2020). The effect of metacognitive strategies implementation on students’ reading comprehension achievement. *International Journal of Instruction*, 13(2), 847–862. <https://doi.org/10.29333/iji.2020.13257a>

- Nguyen, Q. D., Fernandez, N., Karsenti, T., & Charlin, B. (2014). What is reflection? A conceptual analysis of major definitions and a proposal of a five-component model. *Medical Education*, 48(12), 1176–1189. <https://doi.org/10.1111/medu.12583>
- OECD (2019). *What students know and can do. PISA 2018 results / OECD: volume I*. Paris: OECD Publishing. <https://doi.org/10.1787/5f07c754-en>
- Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, 18(1), 22–37. <https://doi.org/10.1080/10888438.2013.827687>
- Rechtsinformationssystem des Bundes (2012). Gesamte Rechtsvorschrift für Lehrpläne der Mittelschulen: BGBl. II Nr. 185/2012. Retrieved from <https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=20007850>
- Reilly, D., Neumann, D. L., & Andrews, G. (2019). Gender differences in reading and writing achievement: Evidence from the National Assessment of Educational Progress (NAEP). *American Psychologist*, 74(4), 445–458. <https://doi.org/10.1037/amp0000356>
- Rumelhart, D. (1977). Toward an interactive model of reading. In Dornic S (Ed.), *Attention and Performance* (Vol. 6, pp. 573–603). Lawrence Erlbaum Associates.
- Scarborough, H. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 97–110). Guilford Press.
- Sigott, G. (2004). *Towards identifying the C-test construct. Language Testing and Evaluation*. Peter Lang.
- Siller, K. (2020). *Predicting item difficulty in a reading test.: A construct identification study of the Austrian 2009 Baseline Reading Test. Language Testing and Evaluation*. Peter Lang. <http://dx.doi.org/10.3726/b16819>
- Siller, K., & Kipman, U. (2018). Cognitive processes as predictors of item difficulty in the Austrian Educational Standards Baseline Test for English Reading at grade 8. In G. Sigott (Ed.), *Language Testing and Evaluation: Vol. 40. Language Testing in Austria. Taking Stock. / Sprachtesten in Österreich: Eine Bestandsaufnahme*. (Vol. 40, pp. 465–488). Peter Lang. <http://dx.doi.org/10.3726/b14338>
- Siller, K., & Kulmhofer, A. (2018). The development of the Austrian Educational Standards Test for English Reading at grade 8. In G. Sigott (Ed.), *Language Testing and Evaluation: Vol. 40. Language Testing in Austria. Taking Stock. / Sprachtesten in Österreich: Eine Bestandsaufnahme*. Peter Lang. <http://dx.doi.org/10.3726/b14338>
- Siller, K., Kulmhofer-Bommer, A., & George, A.C. (2020). Reading for success: Investigating readers' cognitive processes in Austrian EFL reading tests. *PH. Research*. (4), 48–52.
- Singhal, M. (2001). Reading proficiency, reading strategies, metacognitive awareness and L2 readers. *The Reading Matrix*, 1(1).
- Smith, F. (1971). *Understanding reading: A psycholinguistic analysis of reading and learning to read*. Holt, Rinehart & Winston.
- Sparks, R. L., Schneider, E., & Ganschow, L. (2002). Teaching foreign (second) language to at-risk learners. In J. H. Sullivan (Ed.), *Research in second language learning: Vol. 1. Literacy and the second language learner* (Vol. 1, pp. 55–83). Information Age.
- Tummons, J. (2011). 'It sort of feels uncomfortable': problematising the assessment of reflective practice. *Studies in Higher Education*, 36(4), 471–483. <https://doi.org/10.1080/03075071003671794>
- Urquhart, A. H., & Weir, C. J. (1998). *Reading in a second language: Process, product, and practice. Applied linguistics and language study*. Longman.
- Verhoeven, L., Perfetti, C., & Pugh, K. (2019). Cross-linguistic perspectives on second language reading. *Journal of Neurolinguistics*, 50, 1–6. <https://doi.org/10.1016/j.jneuroling.2019.02.001>

- Villacañas de Castro, L. (2016). The pedagogical turn in the research on cross-language transfer: Rethinking the Language Interdependence and the Language Threshold Hypotheses. *Tejuelo (Trujillo)*, (23), 201–227. <https://doi.org/10.17398/1988-8430.23.1.201>
- Williams, E., & Moran, C. (1989). Reading in a foreign language at intermediate and advanced levels with particular reference to English. *Language Teaching*, 22(4), 217–228. <https://doi.org/10.1017/S0261444800014713>

**Klaus Siller** is Senior Lecturer at the Department of Didactics, Instructional and School Development at the Salzburg College of Education Stefan Zweig, Austria. He teaches English (EFL) Methodology and Applied Linguistics. His main publication and research interests are in all areas of foreign language teacher education and particularly in foreign language testing and assessment.

**Andrea Kulmhofer-Bommer** is Senior Researcher at the Austrian Federal Institute of Quality Assurance of the Austrian School System (IQS, Salzburg). She is responsible for the development of the Austrian EFL test for years 7 and 8. Furthermore, she is involved in EFL teacher training. Her main research interests are in language assessment and language teaching methodology with focus on FL reading instruction.

**Ulrike Kipman** is professor at the Department of Educational Sciences and Research at the Salzburg College of Education Stefan Zweig, Austria. Her main research areas are problem solving, gaming, edutainment and decision making. She is an author of many articles and books on improving education in schools.