Kivilehto, M., L. Lahti, T. Pitkänen, E. Pitkäsalo & M. Tervola (toim.) 2024. Tutkimuksellisia siltoja rakentamassa. Vetenskapliga brobyggen. Building bridges through research. AFinLAn vuosikirja 2024. Suomen soveltavan kielitieteen yhdistyksen julkaisuja n:o 81. Jyväskylä. s. 186–210.

Dmitri Leontjev, Ari Huhta & Kristina Koskela
University of Jyväskylä

"I could learn more words": Involving stakeholders in defining L2 English reading constructs in diagnostic and dynamic assessment

Highlights

- L2 English learners' conceptions of reading focus on vocabulary knowledge and are of general kind.
- Teachers and assessment professionals understand reading as integrating different kinds of knowledge.
- Stakeholders' views and theory of L2 reading call for developing reading as an integrated construct.

VERTAISARVIOITU
KOLLEGIALT GRANSKAD
PEER-REVIEWED
www.tsv.fi/tunnus

Abstract

There has been a growing effort to combine two L2 (second/foreign language) assessment frameworks that aim to promote student learning but have different theoretical roots – diagnostic and dynamic assessment. This paper focuses on identifying and operationalising L2 English reading constructs in such a novel assessment framework. We will argue that constructs in this assessment framework should include a view of how they develop and, ultimately, actions that should be taken to guide this development. Focusing on the data from teachers (n=43) and learners (n=725) but also Matriculation Examination item writers and raters (censors; n=22), we propose a procedure for defining and operationalising assessed constructs that combines their theoretical understanding and conceptions, understandings, and actions by these stakeholders. We discuss these data with reference to how constructs and their development are understood in the two assessment frameworks.

Keywords: dynamic assessment, diagnostic tests, foreign language reading instruction, reading comprehension

1 Introduction

Reading is a useful skill for many purposes and contexts ranging from work and education to reading for pleasure. The importance of reading is reflected in the fact that it has been tested in highly influential PISA studies. Not surprisingly, therefore, a vast amount of reading research has been conducted. However, most research concerns reading in the first language (L1), and compared to what is known about L1 reading, our understanding of reading in a second or foreign language (L2) is still rather poorly developed, particularly when it comes to learners' problems in reading and learning to read in L2 (e.g., Alderson 2000; Alderson, Haapakangas et al. 2015).

This article reports on an ongoing study of learners of English in Finnish upper secondary schools (*Lukio* in Finnish) that applies a combination of two approaches to assessment that both support learning, namely diagnostic (DiagA for short) and dynamic assessment (DA). We aim to increase our understanding of reading in English by capitalising on the theoretical understanding of and methods to identify and operationalise reading skills in the two frameworks. We conceptualise reading as a number of interrelated components (or constructs or subskills), which is central to diagnostic L2 assessment, as understanding skills in detail as crucial for understanding learners' strengths and weaknesses so that actionable feedback can be given to them (e.g., Alderson 2005; Lee 2015; Alderson, Haapakangas et al. 2015; Huhta et al. 2024). Diagnostic assessment sheds light on learners' current, independent abilities, but to gain a deeper understanding of learners, we use dynamic assessment that focuses on learner abilities in the process of development (Poehner 2008).

Reading is an important part of learning and teaching English in upper secondary schools in Finland. It is also one of the skills covered in the foreign language tests in

the Matriculation Examination (ME), the final school-leaving examination. However, documents issued by the ME Board do not define reading in L2 in the same way as the rating scale for writing does. The National Curriculum for Upper Secondary Education (FNAE 2019) also lacks a detailed discussion of L2 reading. However, an analysis of ME reading tasks can shed light on how ME censors understand L2 reading and thus provide us with one of the starting points for defining English L2 reading.

In this article, we report on our efforts to define reading in L2 English with reference to constructs. In DiagA, the understanding of constructs derives from psychological and educational research. There, constructs are seen as abstract concepts that describe psychological phenomena such as attitudes, behaviour, and abilities (e.g., Embretson 2010) that have been "inferred from observed behavioral consistencies and that can be meaningfully interpreted" (Chapelle et al. 2008: 3). Reading ability can be understood as a construct, but also the different components or aspects of reading (e.g., understanding the main idea of a text) can be considered (sub-)constructs of reading. What such (sub-)constructs are and how they relate to each other is defined in relevant theories of the skill, as described, for example, by Alderson (2000: 118):

A construct is a psychological concept, which derives from a theory of the ability to be tested. Constructs are the main components of the theory, and the relationship between these components is also specified by the theory.

The understanding of constructs in sociocultural theory, in which dynamic assessment is rooted, somewhat overlaps with how constructs are defined in DiagA. Sociocultural theory emphasises construct development as learners' conceptual development. Furthermore, this development occurs when abstract generalisable constructs descend towards contextual, everyday experiences while those everyday concepts simultaneously move "from the phenomenon upwards towards generalizations" of abstract academic/scientific concepts (Vygotsky 1987: 168; see Poehner 2016 for a detailed metatheoretical discussion). Hence, sociocultural theory views construct development differently from Second Language Acquisition (SLA) research that regards conceptual development as a unidirectional movement towards abstractions and generalisations.

To define and operationalise L2 reading construct and its sub-constructs, we draw on theories developed in SLA and sociocultural theory, and, in particular, research on L1 and L2 reading. We then combined this theoretical understanding with surveys of the key stakeholders in the upper secondary context, namely teachers of English and students in *Lukio*, exploring also the ME item writers' and raters' (hereinafter, censors) views.

First, we provide a brief overview of the two approaches to L2 assessment that our research relies on, focusing on how constructs are defined and operationalised in them and what the role of teachers and learners is in this process. We then report on how different stakeholders conceptualise L2 English reading and how they think reading in English could be developed.

2 L2 Diagnostic and Dynamic assessment and constructs identification and operationalisation

2.1 Diagnostic assessment: Focusing on constructs

Diagnostic assessment in L2 aims to identify learners' strengths and particularly weaknesses and, ultimately, understand the reasons for underlying weaknesses so that meaningful action can be taken to address these weaknesses. Serious attention to diagnostic assessment in applied linguistics began in the late 1990s with the design of DIALANG, a large-scale, multilingual assessment system (Alderson 2005).

Current thinking on L2 diagnostic assessment emphasises the importance of understanding DiagA as a process or a cycle consisting of phases necessary for useful diagnosis (Alderson, Brunfaut et al. 2015; Alderson, Haapakangas et al. 2015; Huhta et al. 2024). Defining the constructs is the crucial starting point of the cycle, followed by the design (or selection) of assessment instruments or procedures that tap the intended constructs. The assessment itself can happen as self-assessment by the learners, but more typically by the teacher or some other external agent. The assessment results then need to be interpreted and turned into feedback to the learner.

Two points emerge from this brief outline of L2 DiagA as particularly relevant for developing diagnosis: (a) L2 diagnosis should be firmly grounded in a relevant theory, and (b) teachers and learners are important stakeholders and should be involved in the assessment development process.

For diagnosing L2, the general theoretical basis is Second Language Acquisition research (Alderson 2005) and its different strands. A fairly robust SLA literature exists on L2 reading (e.g., Alderson 2000; Alderson, Haapakangas et al. 2015; Grabe & Yamashita 2022), but in comparison with the vast field of L1 reading research, work on L2 reading is more limited. Therefore, L2 researchers regularly draw on L1 research (Alderson 2000; Alderson, Haapakangas et al. 2015).

Research literature presents several lists of reading constructs, none of them comprehensive and agreed on by everybody. As a practical solution, we focus on the constructs that are most common in L2 reading and assessment research. An important source is Grabe and Yamashita's (2022) recent synthesis of L2 reading research. It lists major component abilities for reading comprehension, such as vocabulary knowledge, syntactic knowledge, awareness of text structure and discourse organisation, main idea comprehension, recall of relevant details, inference, and evaluation and critical reading.

Research shows that reading in L1 or L2 is a complex process in which both bottom-up (e.g., word recognition and syntactic knowledge) and top-down (e.g., background knowledge) information contribute to understanding, and different skills can compensate for each other (Grabe & Yamashita 2022). While recognising this complex

interplay between different types of knowledge in reading, for diagnosing reading, we seek to define a number of reading constructs and analyse reading tasks with reference to these constructs so that learners' L2 reading can be improved by using exercises that target them. Another key research finding is that two types of models are involved in reading: a model of text meaning and a situational model (e.g., Kintsch 1998). The former refers to what the writer of the text wants to convey, and the latter relates to the personal meaning(s) of the text to the reader. Such models suggest that there are more global understandings of the texts which may or may not be shared by the writer and reader. These understandings can also arise, at least in part, as a result of the above-mentioned interplay of different reading constructs.

A weak point in the typical diagnostic assessment process is the link between feedback and action that should follow from it. Since DiagA focuses on assessment, literature on DiagA has not yet addressed this question in detail; thus, it is not fully clear what specific action should follow particular feedback, although the different types of feedback discussed in the literature provide some broad guidelines (e.g. Hattie & Timperley 2007; Huhta et al. 2024). Dynamic assessment can help address this problem by first integrating teaching/learning in assessment sessions and, second, by expanding the scope of the assessed constructs from learners' current abilities to those that are only emerging. Furthermore, the emphasis on the development of the DA approach to assessment may also help L2 readers construct textual and situational models that are typical of the reading process. DiagA's typical focus on sub-constructs, while useful for deciding what subsequent training should focus on, has the potential danger of not adequately covering the more holistic aspects of reading.

Next, we describe the development of the reading tests in DIALANG as an example of how diagnostic L2 reading assessments are designed. DIALANG is an online diagnostic assessment and feedback system designed in a large European project between 1997-2004 (Alderson 2005; Alderson, Haapakangas et al. 2015). It tests reading, listening, writing, vocabulary, and structures in 14 languages. Overall, the test design process in DIALANG was similar to the careful, multi-stage development used in high-stakes examinations. In other words, an assessment framework and detailed test specifications were created for each skill to guide the item writing. An important aspect of this stage was defining the sub-skills to be targeted, which for reading were understanding the main idea, understanding details, and making inferences.

The test items were then reviewed and selected for piloting online, after which items were chosen for the operational tests based on statistical analyses. Because DIALANG uses the 6-point scale of the Common European Framework of Reference (CEFR; Council of Europe 2001) to report the overall results, standard-setting procedures were also employed to link the scores with the CEFR levels (Alderson 2005). However, the three sub-skill scores are reported simply as the number of correct and incorrect responses, and the learner has to figure out their strengths and weaknesses

by comparing the proportions of correct answers for the different sub-skills. This is probably the most common way to interpret subskill scores in other diagnostic L2 tests.

Learners and teachers can play an important role in the development of L2 diagnosis; however, this varies. In DIALANG, learners' role was limited to providing response and background data in the pilot testing phase (Alderson 2005). Teachers' main role has been to help students interpret DIALANG feedback when the system is used in educational settings. A somewhat similar role for teachers is envisaged in more recent conceptual treatments of ideal L2 diagnosis (e.g., Alderson, Brunfaut et al. 2015; Harding et al. 2015), where the teacher is envisaged to conduct an initial assessment of learner difficulties and to decide on appropriate action. However, they recommend that teachers use professionally developed instruments for the actual diagnosis (see also Huhta et al. 2024 for a discussion of the challenges in diagnostic and formative assessment in the classroom).

Even if learners and teachers have had very limited input in the design of diagnostic L2 tests so far, they have a key role in any validation study of such tests. For example, the learners using DIALANG have been investigated for their views on the usefulness of the feedback from the system (Huhta 2010). It should be noted that there are calls in the more general L2 assessment literature for the increasing involvement of learners also in the test design process (see Jin 2023, for an overview).

2.2 Dynamic assessment: Focusing on development

Dynamic assessment (DA) is rooted in Vygotskian sociocultural theory, and it brings to light learner abilities that are in the process of development. Vygotsky (1987) reasoned that common assessments focus on abilities that have fully matured and that learners can demonstrate independently during the assessment. However, by giving learners assistance during assessment, more information can be obtained. Such support sheds light on what learners can do with help from others. Furthermore, learners' responsiveness to support when they encounter difficulties during assessment varies, which reveals how close learners are to independent functioning. Vygotsky (1987, 1998) discussed this range of emerging abilities with reference to the concept of Zone of Proximal Development (ZPD), arguing for its significance, as it is the abilities in the process of development that are most susceptible to instruction (see Poehner & Leontjev 2020). Learners' ZPD, then, becomes the object of diagnosis during DA.

As Poehner and Leontjev (2020) discussed, it is not only the amount of external support, known as mediation, which is important during DA, but also its quality and the kind of support given to learners to develop their abilities. Indeed, while the amount of support that learners receive during DA gives an idea of how close they are to independent functioning, the quality of this support allows for a particular ZPD to emerge. This also allows the expert mediator to develop learners' abilities in the

direction intended by the mediator. How learner development is directed is based on the mediator's emergent understanding of learners' ZPD. DA then combines assessment and teaching into one development-oriented process, making them interdependent (Infante & Poehner 2019).

The focus of mediation on learner abilities in the process of maturing is particularly relevant to our argument for construct identification and operationalisation in a framework that merges DiagA and DA. In sociocultural theory, individuals' development is understood with reference to the emergence of, in Vygotsky's (1998: 189) words, "qualitative neoformations", the construction of novel mental functions leading to the reorganisation of the whole system of consciousness (Veresov & Mok 2018). In other words, it has to do with individuals' conceptual development as novel understandings, knowledge, and means for mediating thinking processes are internalised. It is important to keep in mind that, in DA, the understanding of constructs is holistic. The focus is not on a number of narrower aspects constructing a whole, each aspect possible to study separately. Rather, it is on the whole from the outset, defined with reference to an individual's history and a network of relations (Poehner 2016). From this angle, a view of reading such as the integration of several kinds of knowledge unique to individual readers, as we elaborated earlier, seems to be more in line with how constructs are understood in DA.

The understanding of development in sociocultural theory expands the role of theory in construct identification in DiagA and the roles of stakeholders, particularly teachers and learners, in this process. Namely, DA is informed by a coherent theory of individuals' development, compelling us to understand learner performance during assessment as an indication of their conceptual development, and to intervene in this process to understand it more fully with the goal of transforming learner thinking and functioning. In L2 DA, this requires understanding what challenges learners have with a particular construct, what their conceptions of it are, and what partial understandings they may have that could be built upon. Hence, learners and teachers should be involved in developing DA, which inspired our study to involve learners as key stakeholders as we need to understand how they understand L2 reading.

Many L2 DA applications have been taken in classroom settings in collaboration with teachers (e.g., Poehner & Wang 2021), meaning that the targets of these DA implementations were informed by curricula rather than a theory of language proficiency. When L2 dynamic assessments are designed to stretch beyond specific classrooms, learners and teachers are still involved in their development. Several studies focusing on L2 DA of reading are useful to mention here. Teo (2012) reported on action research where a computerised DA procedure was designed to address specific challenges with inferential reading that the learners the author taught had. The author used a pretest-intervention-posttest design in which the pretest was used to understand the learners' specific challenges with inferential reading in L2 English which were then addressed in a computerised DA intervention focusing on (a) the

concept of inferential reading and (b) strategies, and importantly, their meanings for inferential reading that learners could use. The most general and implicit mediation in the study first provided a simple explanation for inferencing as reading between the lines. It continued by explaining what this means for the reading process: "inferential question cannot be answered by looking at the text itself" and "you will need to use the information stated in the passage to infer what is not stated" before requiring the learners to think about the main idea of the passage. Only then was a strategy introduced that the learners could use, focusing on keywords and parts of the text. The author found that the procedure was, indeed, beneficial for developing not only learners' performance but also their comprehension strategies (their metacognition), as was evidenced in the learners' reflection in learning portfolios.

Poehner et al.'s (2015) study involved learners in the design process of a computerised DA of L2 Chinese reading (and listening) in a somewhat different way from Teo (2012). First, the tests were piloted in a non-DA format with 28 learners to identify whether the intended constructs were reflected in the tests. Next, the assessors engaged with six learners in dynamic assessment sessions with flexible probing and guiding to reveal their partial understanding and how it can be supported and promoted to develop their reading comprehension ability.

However, L2 DA researchers usually work with teachers in small-scale projects, basing their understanding of constructs they focus on and learners' areas of struggle on the teachers' understanding (see Poehner & Leontjev 2020). Regarding constructs and their development, in DA as in DiagA, teachers and learners are essential stakeholders involved in the process of their identification and operationalisation.

3 Methodology

In this study, we report on our early work in bringing together L2 DiagA and DA frameworks, proposing how understandings of constructs in them and methods regarding construct identification, conceptualisation and operationalisation can be combined. We focus on teachers and learners in this process while also discussing the role of L2 assessment experts. Our construct work in the dynamic-diagnostic assessment framework is strongly based on combining the theoretical understanding of L2 reading and its development, and conceptions and practices of mainly two groups of stakeholders: teachers and students in the upper secondary school (*Lukio*).

3.1 Research questions

The research questions that guided our work were the following:

RQ1. How do different stakeholders conceptualise reading in English in the Finnish upper secondary school?
RQ2. How can Finnish upper secondary school students' reading in English be developed according to teachers and learners?

When engaging with the stakeholder groups, we used methods typical of the two assessment frameworks, keeping also in mind how constructs are understood in both frameworks. This included, for example, investigating how learners understand L2 reading construct as a part of its operationalisation. However, it should be noted that the data we discuss in this paper are only a part of the construct work in our effort to bring DiagA and DA together.

3.2 Overall design and steps in defining reading constructs

First, we defined reading constructs by combining theoretical and empirical information about reading in L2. The theoretical literature on L2 reading and assessment of L2 reading (summarised in, e.g., Alderson 2000; Alderson, Haapakangas et al. 2015; Grabe & Yamashita 2022) provided us with a large number of potential constructs at different levels of granularity, against which we could interpret the empirical data collected from the three stakeholder groups using online survey questionnaires. This first stage comprised procedures typical for designing diagnostic assessments, that is, the aim was to develop L2 reading tasks for English that tap specific reading constructs and yield information about the learners' strengths and weaknesses on those constructs. This article focuses on this stage of the project and reports on the empirical findings collected through the surveys. The following stage, based on dynamic assessment approaches and different empirical data, is underway but is not covered here, with the exception of some references to it in Section 5.

Early on in the project, the decision was made to use only existing, high-quality reading tasks because the empirical piloting of reading tests is labourious and time-consuming. Therefore, past ME English reading items and items from DIALANG English reading tests were selected. Additionally, using ME tasks in the research is likely to motivate students since past examinations are regularly used in schools as a way to prepare for the examination.

The three types of online questionnaires were administered in a particular sequence because some of the key content in the latter two were based on the responses given to the first. The first questionnaire targeted the ME censors who were requested to describe in their own words what they thought selected items from past English ME tests assessed and what challenges the students might encounter in them. These

open-ended answers were then categorised into a dozen classes (i.e., constructs) that were used in the subsequent questionnaires for the students and teachers of English as the categories in multiple-selection questions about the challenges the students have with the selected reading items.

3.3 Participants and data

The research data reported here come from three questionnaires directed to three large stakeholder groups – (a) ME censors, (b) upper-secondary school teachers, and (c) students. The following Figure 1 outlines the participants and the data in this study.

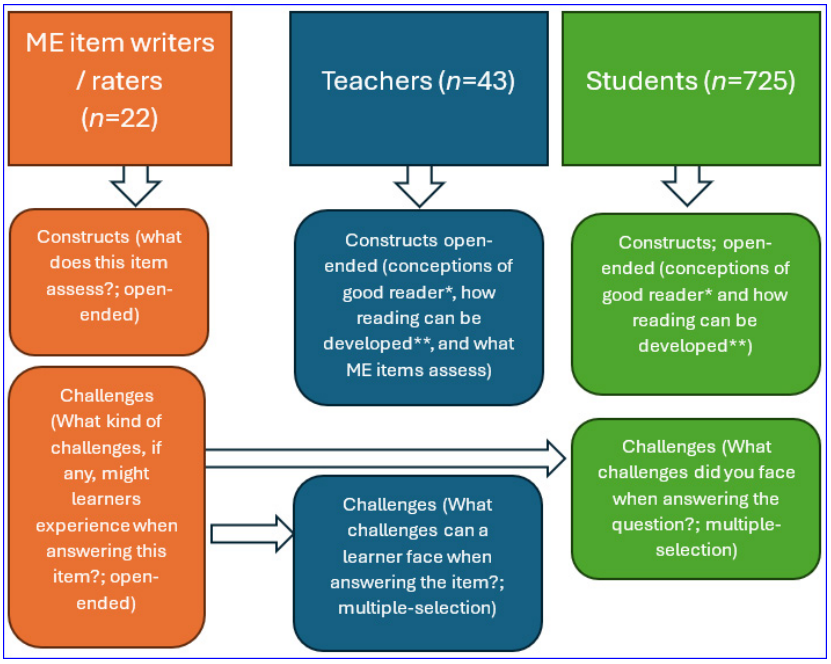

FIGURE 1. Participants and data

*The question was slightly different for teachers and learners. For the teachers: What should a *Lukio* student be able to do/know to be a good reader in English?; for the students: Think of someone who you think is good at reading in English. What do you think such a person can do?

** The question was slightly different for teachers and learners. For the teachers: What do you do to help your learners (a) who are struggling in their learning process and (b) those learners who are fairly good at reading become better readers in English?; for the students: What kind of things do you think you should develop to read better in English?

The three questionnaires had somewhat different foci although they shared certain core questions. To contextualise the gathering of stakeholder views about L2 (English) reading constructs and, thus, make replying easier, all questionnaires contained examples of past ME reading texts with their related questions. The respondents were asked to try to answer each question and tell either what the particular item (i.e., text + question) tested in their view or what challenges they (for students) or a typical student (for ME representatives and teachers) encountered when taking the item. We consider both types of questions (what constructs?; what challenges?) to shed light on the respondents' views about reading constructs, albeit from somewhat different perspectives.

The ME censors' questionnaire contained only questions targeting specific ME items whereas the teachers' and students' questionnaires included fewer ME items but covered several aspects of learning, teaching and assessing reading in English. Teachers' and students' questionnaires included both multiple-selection questions about challenges - these were always connected with concrete ME reading items - and an open-ended question that allowed them to describe a good reader in English in their own words.

As explained, the ME censors' questionnaire was administered first since its open-ended questions about reading constructs and challenges were categorised into types to make answering them quicker for the teachers and students. The ME English item writers responsible for designing the English tests (n=5) were asked to respond to a total of twenty English ME reading items. A different group of ME language item writers (n=5) involved in reviewing the English ME and English ME raters (n=12) answered ten ME items out of twenty that the English item writers took. In both teachers' and learners' questionnaires, six ME items out of the initial pool of 20 items were included by randomly assigning two ME items to each participant.

In both questionnaires for teachers (n=43) and the students (n=725), the respondents were asked open-ended questions about their perceptions of what someone should be able to do/know to be a good reader in English, what kinds of challenges or points for development *Lukio* students have in reading in English, and what the *Lukio* teachers do or could do to help learners address the challenges.

All questionnaires were piloted among colleagues specialising in survey research and language assessment. The teacher questionnaire was also piloted among several English teachers. The learner questionnaire was piloted in part with a group of MA students participating in a language assessment course at a university in Finland. The questionnaire items we focus on in this paper are presented in Appendix 1.

The open-ended questionnaire data were analysed using content analysis. Both *a priori* coding and grounded coding were used for these data. The three researchers coded all the questionnaire items, dividing them so that each coded a part of the sample with some overlap. Everyone, hence, coded about two-thirds of the responses. The *a priori* coding was based on our close engagement with SLA research on L2 reading constructs. The coding started with ME censor data. The same coding scheme was later applied to the other two questionnaire data sets. The grounded coding and adjustments to the *a priori* coding scheme were discussed in weekly data sessions.

4 Results

The results will be presented in two sections, each focusing on one research question. We first report on different stakeholders' conceptions of L2 English reading and the challenges that learners can encounter.

4.1 The construct of L2 English reading: Stakeholders' conceptions and perspectives

One step in the inquiry into the L2 English reading construct was getting an overview of general decontextualised conceptions of L2 English reading by teachers and learners. These data come from two slightly differently formulated open-ended questions.

The learners' responses are summarised in Figure 2.

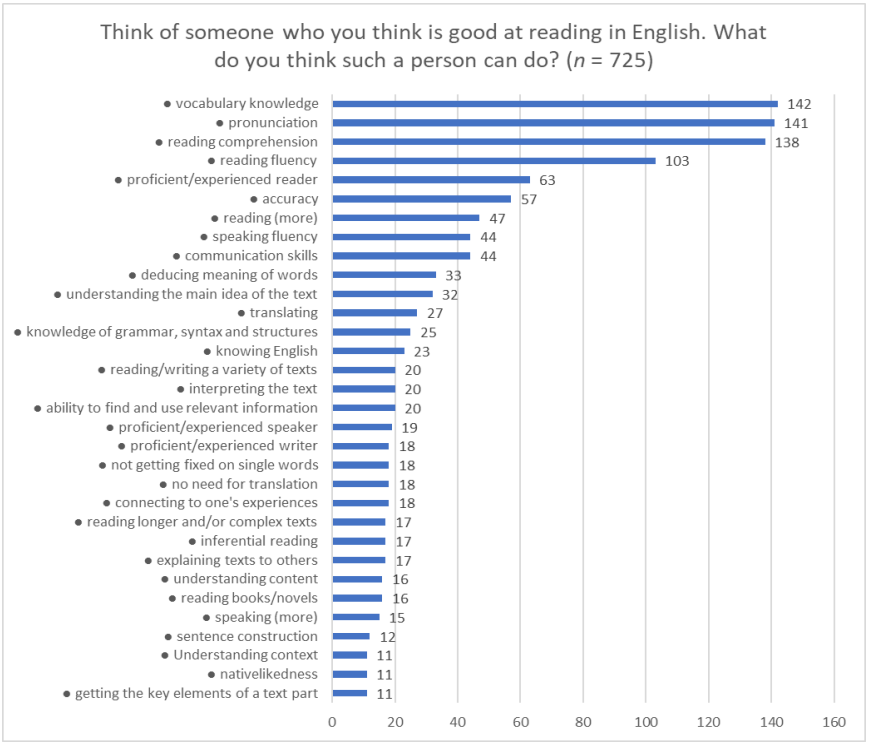

FIGURE 2. What learners (n=725) indicated when asked what a person who is good at reading in English can do.

The views of reading by most learners appeared to be rather general, involving mainly vocabulary, fluency and, interestingly, accuracy. It was, indeed, rather unexpected that the learners mentioned accuracy and pronunciation as notable characteristics of a proficient reader. A reason for this might be that they imagined a good reader in English to also be a proficient user of the language, with pronunciation and accuracy being parts of their conception of L2 English proficiency. It is also possible that reading textbook texts aloud is one of the classroom activities in some schools. Admittedly, learners mentioned more specific reading purposes, too, such as inferential reading, reading for the main idea, being able to explain texts for others, or indeed, bringing other kinds of knowledge about language, e.g., grammar. However, only a few mentioned these, and mostly they responded in a rather general way.

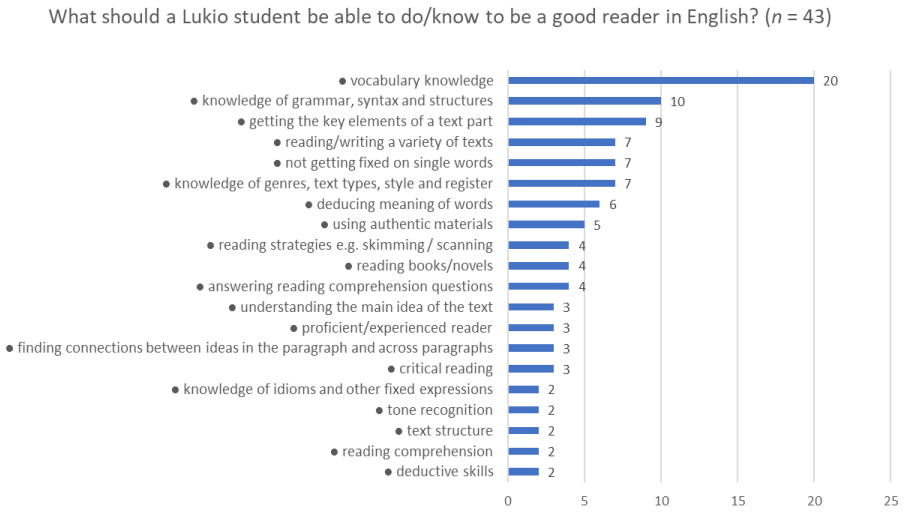

FIGURE 3. What teachers (n=43) indicated a *Lukio* student who is a good reader should be able to do.

The teachers' responses (Figure 3) were often more specific. They mentioned various reading purposes, reading strategies, and kinds of knowledge about language that a *Lukio* student who is a good reader should have.

In addition to the open-ended questions asking the teachers and students to describe good readers in English, we solicited the respondents' views about reading with reference to concrete reading tasks. Each respondent was asked to take two English reading comprehension items selected from past Matriculation Examinations and indicate the challenges they encountered when taking a particular item by se-

lecting them from a list of challenges shown to them. This list was designed based on the content analysis of the ME censors' responses to open-ended questions about the same items.

The responses to the questions regarding the challenges related to specific ME reading items shed further light on their conceptions of the constructs involved in reading in English. The responses indicated both differences and similarities across the three groups, with teachers' and ME item censors' views being rather similar, whereas the students' views differed from theirs. Due to lack of space, we will focus on two questionnaire items (Appendix 1). The first related to the text *Electric Car* used in the Autumn 2022 advanced-syllabus English exam. The item *What is mentioned as a difference between GM and Tesla?* was a multiple-choice question, which, according to our understanding and that expressed by most ME censors, taps the construct of reading for specific detail but also learners' vocabulary knowledge and their ability to deduce the meaning of unknown words from context.

The first figure in Appendix 2 summarises the proportion of (a) ME censors, (b) teachers, and (c) learners who mentioned a particular challenge. We repeat that the frequencies/proportions were calculated differently for the ME censors compared to the other two groups. In the ME group, the frequencies come from the content analysis, which served as the basis for the multiple-selection item used for the other two groups. The formulation of the specific question was the same for the teachers and ME censors (*What challenges can a learner face when answering the item?*) but slightly different for the learners (*What challenges did you face when answering the question?*). Our intention was to elicit the groups' views on what challenges *learners* might have, which meant that we asked the teachers to consider the matter from their students' perspective rather than what they themselves found difficult when taking the ME items, although that personal experience likely affected their responses to some extent.

The challenges indicated by the teachers and ME censors were rather similar: both groups generally selected a small number of similar challenges. One obvious difference was that all teachers marked understanding specific details as a challenge, but none of the ME censors did so. In our opinion, this may reflect differences in how they understood the constructs being assessed. We assume that this was because ME censors wished to be more precise in identifying learners' challenges with the item, focusing on what could become an obstacle when learners responded to it. Regarding the students, they usually selected the same challenges as the teachers and the ME censors, but they also selected many other points in the list. This suggests that their challenges may be different from those assumed by the teachers and ME censors or that they may, in fact, be uncertain about the challenges they faced, which some of the learners' responses, such as "answering in Finnish", imply.

A somewhat different picture emerges from the analysis of the second ME item that illustrates our construct work – a variant of a multiple-choice task, where parts of the text *Climate Crisis* were replaced with gaps, each with several options in a drop-

down menu. This item appeared in the examination as the previous item (Appendix 1). The same analysis was used for this second figure in Appendix 2 as for the first figure in this appendix.

In this item, unlike in the first ME item, the teachers' and learners' reports are rather similar in that they selected most of the challenges listed, whereas the ME group generally chose only a few challenges The notable challenge on which all three groups agreed was combining information from different parts of the text, which is understandable as considering the text that immediately precedes and follows the gap is required to answer this item. The large number of challenges selected by both the teachers and the learners suggests that this kind of item may require combining different kinds of knowledge, including multiliteracy, background knowledge, and inferential reading.

In general, juxtaposing the teachers' and the learners' general understanding of the reading construct with their thinking about specific ME items allows for seeing a notable difference. Teachers' and learners' responses to the decontextualised open-ended questions suggested that both groups (particularly learners) had either a rather general, even vague, idea of what L2 reading is about or that certain features such as vocabulary were particularly important in L2 reading. However, the fact that they typically selected several features as challenges when the question was presented contextualised with a concrete L2 reading item, suggests that they considered these ME items requiring integration, or at least, the use of several kinds of knowledge.

4.2 Developing learners' L2 reading ability

In this section, we discuss data that should provide a more solid basis for developing learners' conceptions of L2 reading, which is at the core of DA. The learner data come from a questionnaire item that asked the learners to elaborate on what they think they need to develop to become better readers in English. Figure 4 summarises their responses.

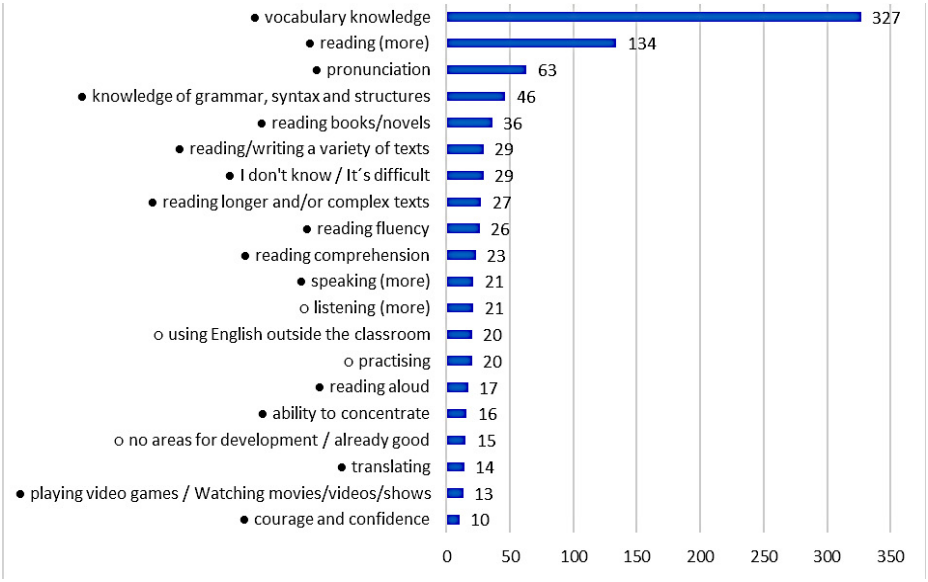

FIGURE 4. Learners' (n=669) responses to the item *What do you think you need to develop to read better in English*.

Figure 4 shows that the learners' views of how they could improve their reading are rather general and largely repeat their conceptions of a good reader. However, it gives a basis for engaging with the learners to develop their understanding of their own developmental process.

To investigate the teachers' understanding of how learners' reading can be developed, we used two complementary ways, both using the coded teacher responses to the respective questionnaire items. The first focuses on the topics that emerged in the teachers' responses to a general question as to how they develop learners' reading (see Figure 5). We excluded topics that emerged only once from the figure.

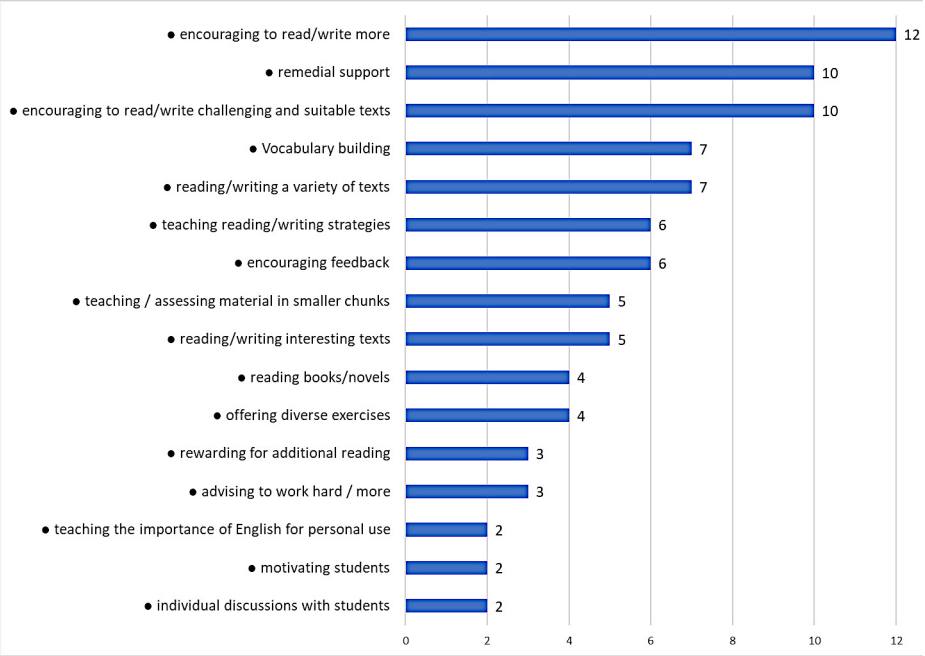

FIGURE 5. What teachers reported can be done to develop learners' reading comprehension ability (n=43).

As far as specific constructs are concerned, vocabulary emerged most often in our sample. The rest of the topics largely fall under encouraging and motivating learners and the versatility of texts and tasks used to develop reading ability. Teaching reading strategies, including scaffolding reading tasks by approving them in smaller chunks, are too, worth mentioning. The points mentioned only by individual teachers included (a) more specific reading strategies, not getting fixed on single words, or focusing on text structure, (b) a teaching strategy of simplifying texts, (c) constructs of genre, text types, styles, and registers, and (d) activities such as practising reading comprehension questions, quizzes, and pairwork. From this general question, it emerges that the ways teachers and learners thought about developing reading were somewhat similar.

To complement these findings, we next illustrate our analysis of what the teachers reported with regard to the item for the *Climate Crisis* text discussed earlier (Appendix 1). We limit our discussion to this item, as both ME items give quite a similar picture of how the teachers reported they help their students develop the skills needed to answer the item. Figure 6 outlines the themes that the teachers mentioned when responding to the item, which we further grouped into larger categories of constructs,

challenges, strategies, and activities. As sometimes it was difficult to decide whether teachers described strategies or classroom activities, we coded parts of the statements as both (orange codes in Figure 6).

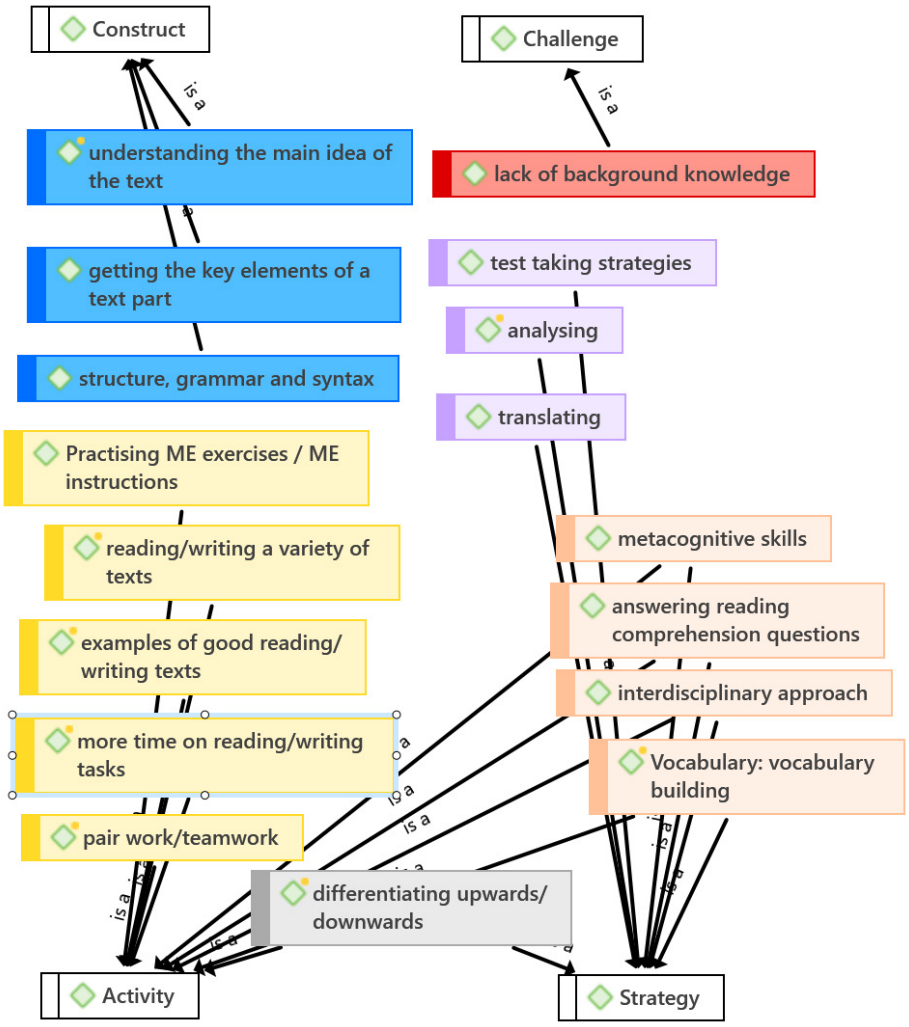

FIGURE 6. What teachers (n=11) reported can be done to develop learners' reading comprehension ability needed to solve ME items like the one to the *Climate Crisis* item.

Several conclusions can be drawn from teachers' responses to the item in Figure 6. First, the teachers brought in their reflection both particular constructs and challenges that learners may have, which partially covered the same constructs that the teachers

thought the item assessed (see Appendix 2). More relevant to our purposes, however, were the kinds of activities the teachers would use and the strategies they would equip their learners with. Some of these were quite understandable given that these were items from the ME: teachers mentioned certain test-taking strategies (mostly elimination), and practising ME and other reading comprehension tasks. However, several strategies and activities seem to refer to reading more generally. These included analysing the text, translating (i.e., using L1 as a resource to understand the text), interdisciplinary approach (related to multiliteracy and disciplinary knowledge), and developing learners' metacognitive skills (learners' awareness of their reading processes and knowledge about the language they apply to understand the text).

5 Discussion and conclusion

The present study was part of our ongoing work in conceptualising, identifying, and operationalising L2 English reading constructs in the framework that combines the two approaches to assessment that support learning—diagnostic and dynamic assessment. Both approaches involve an understanding of how learner abilities can be developed and are complementary in that while DiagA involves a detailed and fine-grained understanding of the assessed constructs usually stemming from L2 research, DA, in turn, is strongly based on a coherent theory informing the understanding of how learner development happens.

The combined data from teachers and learners particularly regarding our analysis of their engagement with the single ME L2 reading items points towards a view of reading where readers integrate different kinds of knowledge about language to comprehend texts (see Grabe & Yamashita 2022). Our analysis also allowed us to target our assessment to such constructs as specific details and main ideas, but also deducing vocabulary from the context. From the diagnostic assessment perspective, these data are very useful. We also obtained a picture of how teachers and students understood L2 reading. Particularly the students' conceptualisation was rather general, focusing on vocabulary, or discussing it with reference to such constructs as fluency and accuracy. Only a few learners mentioned such constructs as grammatical and syntactic knowledge, inferential reading, or positioning the text within the larger context, for example.

Teachers were more specific than learners in that their concept of L2 English reading was, understandably, more detailed. However, neither teachers nor learners mentioned the integration of several kinds of possibly compensatory types of knowledge about language. This opens up interesting opportunities for dynamic assessment. In it, the learners' concept of reading can serve as a starting point for developing this understanding of reading through mediation building on the emerging under-

standing and types of knowledge learners possess. That is, the development of L2 reading comprehension for Finnish L2 English learners should involve the expansion of their understanding of reading as largely vocabulary knowledge to a construct that requires integration of different kinds of knowledge, which can be different for different learners, depending on their strengths and weaknesses in English.

Teachers' practices in developing L2 reading abilities of their learners, in which their responses to the corresponding survey question used to answer the second research question can illuminate this process. Before moving our discussion to teachers, we note that, for learners, the areas for development and ways for developing these largely correspond to their conceptions of reading, including developing their vocabulary and reading more and more broadly.

For teachers, in addition to helping learners develop their vocabulary and, in general, creating opportunities for practising reading, equipping learners with various strategies was important, particularly regarding their engagement with specific ME items. At the same time, only one teacher mentioned developing their learners' metacognition (Figure 6). However, being conscious of one's own reading processes can be seen as essential for developing learner abilities according to sociocultural theory, which is the basis for DA.

Our approach to construct identification brought together (a) a theoretical understanding of constructs and (b) methods involving teachers and learners in construct identification used in DiagA and DA. This allowed us to identify areas for learner development in a detailed and theoretically informed way and to develop our understanding of the starting point for learner development in L2 reading as well as propose how the development can be guided. In other words, DiagA, especially its roots in SLA and L2 research, gave us a basis for identifying the construct of L2 reading and its sub-constructs, whereas the theory behind DA further informed the understanding of these constructs and their development. Both, in turn, informed the way we engaged teachers and learners in this study.

Based on our findings, we propose that the guidance/mediation during the dynamic-diagnostic assessment of L2 English reading should (a) develop learners' understanding of how they can mobilise their knowledge about language in an integrated way, (b) include strategies and (c) promote learner conceptual understanding of reading and metacognition of their own reading processes.

To pave the way for designing mediation for L2 reading, we already engaged in applying the findings in further construct work, including exploring the reading processes of expert L2 readers, and piloting the mediation with several *Lukio* students. This will allow us to further refine our understanding of the L2 reading construct (and its sub-constructs) in assessment that brings together DiagA and DA, including its theoretical, metatheoretical, and methodological bases and traditions. We admit that one challenge that is yet to be addressed is the commensurability of the theoretical

bases of DiagA and DA, which we do not discuss in this paper. The study we report in this paper is one step in this process.

As far as limitations are concerned, our sample size of the *Lukio* teachers was rather small and thus cannot be seen as representative of all the *Lukio* English teachers in Finland. The reader could also argue that our sample of ME censors is also small, but it is, in fact, rather small to start with, meaning that we were able to recruit most of the item writers in particular.

Beyond identifying the reading constructs in a dynamic-diagnostic assessment, our study allows for proposing a general approach for identifying such constructs, involving both theory and data from various stakeholders. These latter data, we propose, should focus both on conceptions of the particular (sub-)construct by learners (and teachers) and ways that it can be developed. As we hopefully demonstrated, both should be useful for such construct work.

Literature

Alderson, J. C. 2000. *Assessing Reading*. Cambridge: Cambridge University Press.
Alderson, J. C. 2005. *Diagnosing foreign language proficiency: the interface between learning and assessment*. London: Continuum.
Alderson, J. C., E.-L. Haapakangas, A. Huhta, L. Nieminen, & R. Ullakonoja 2015. *The diagnosis of reading in a second or foreign language*. New York, NY: Routledge.
Alderson, J. C., T. Brunfaut, & L. Harding 2015. Towards a theory of diagnosis in second and foreign language assessment: insights from professional practice across diverse fields. *Applied Linguistics, 36* (2), 236–260. https://doi.org/10.1093/applin/amt046
Chapelle, C., M. Enright & J. Jamieson 2008. Test score interpretation and use. In C. Chapelle, M. Enright & J. Jamieson (eds) *Building a validity argument for the Test of English as a Foreign Language*. New York, NY: Routledge.
Council of Europe 2001. *Common European framework of reference for languages: learning, teaching and assessment*. Cambridge: Cambridge University Press.
Embretson, S. (ed.) 2010. *Measuring psychological constructs*. Washington, DC: American Psychological Association.
FNAE 2019. *Lukion opetussuunnitelman perusteet*. Helsinki: Opetushallitus.
Grabe, W. & J. Yamashita 2022. *Reading in a second language: moving from theory to practice.* Cambridge: Cambridge University Press.
Harding, L., J. C. Alderson, & T. Brunfaut 2015. Diagnostic assessment of reading and listening in a second or foreign language: elaborating on diagnostic principles. *Language Testing, 32* (3), 317–336. https://doi.org/10.1177/0265532214564505
Hattie, J. & H. Timperley 2007. The power of feedback. *Review of Educational Research, 77* (1), 81–112. https://doi.org/10.3102/003465430298487
Huhta, A. 2010. *Innovations in diagnostic assessment and feedback: an analysis of the usefulness of the DIALANG language assessment system.* An unpublished doctoral dissertation. University of Jyväskylä.
Huhta, A., C. Harsch, D. Leontjev, & L. Nieminen 2024. *The diagnosis of writing in a second or foreign Language*. New York, NY: Routledge.
Infante, P., & M. E. Poehner 2019. Realizing the ZPD in second language education: the complementary contributions of dynamic assessment and mediated development. *Language and Sociocultural Theory, 6* (Special Issue 1), 63–91. https://doi.org/10.1558/lst.38916
Jin, Y. 2023. Test-taker insights for language assessment policies and practices. *Language Testing*, *40* (1), 193–203. https://doi.org/10.1177/02655322221117136
Kintsch, W. 1998. *Comprehension: a framework for cognition*. New York, NY: Cambridge University Press.
Lee, Y.-W. 2015. Diagnosing diagnostic language assessment. *Language Testing*, *32* (3), 299–316. https://doi.org/10.1177/0265532214565387
Poehner, M. E. 2008. *Dynamic assessment: a Vygotskian approach to understanding and promoting L2 development*. Boston, MA: Springer.
Poehner, M. E. 2016. Sociocultural theory and the dialectical-materialist approach to L2 development. *Language and Sociocultural Theory, 3* (2), 133–152. https://doi.org/10.1558/LST.V3I2.32869
Poehner, M. E. & D. Leontjev 2020. To correct or to cooperate: mediational processes and L2 development. *Language Teaching Research*, *24* (3), 295–316. https://doi.org/10.1177/1362168818783212

Poehner, M. E., J. Zhang, & X. Lu 2015. Computerized dynamic assessment (C-DA): diagnosing L2 development according to learner responsiveness to mediation. *Language Testing, 32* (3), 337–357. https://doi.org/10.1177/0265532214560390

Poehner, M. E. & Z. Wang 2021. Dynamic assessment and second language development. *Language Teaching*, *54* (4), 472–490. https://doi.org/10.1017/s0261444820000555

Teo, A. 2012. Promoting EFL students' inferential reading skills through computerized dynamic assessment. *Language Learning & Technology, 16* (3), 10–20. http://dx.doi.org/10125/44292

Veresov, N. & N. Mok 2018. Understanding Development Through the Perezhivanie of Learning. In J. P. Lantolf, M. E. Poehner, M. Swain (eds) *The Routledge Handbook of Sociocultural Theory and Second Language Development*. New York: Routledge, 89–101.

Vygotsky L. S. 1987. *The collected works of L. S. Vygotsky: Volume 1: problems of general psychology, including the volume thinking and speech.* R. W. Rieber & A. S. Carton (eds). New York, NY: Plenum.

Vygotsky L. S. 1998. The problem of age. In R. W. Rieber (ed.) *The collected works of L. S. Vygotsky: Volume 5: child psychology*. New York, NY: Plenum, 187–206.

Yle 2022. *10. Electric Car.* Matriculation examination. Advanced-syllabus English exam. Fall 2022. https://yle.fi/plus/abitreenit/2022/Syksy/2022-09-16_EA_fi/attachments/index.html#10.A and https://yle.fi/plus/abitreenit/2022/Syksy/2022-09-16_EA_fi/index.html#10 [Accessed 2 July 2024.]

Yle 2022. *12. Climate Crisis.* Matriculation examination. Advanced-syllabus English exam. Fall 2022. https://yle.fi/plus/abitreenit/2022/Syksy/2022-09-16_EA_fi/index.html#12 [Accessed 2 July 2024.]

APPENDIX 1

The two ME items discussed in the study (links).

As the Matriculation Examination materials are in open access, the links below are provided.

1) *Electric Car* text and item (10.3 *What is mentioned as a difference between GM and Tesla?*):
   - » https://yle.fi/plus/abitreenit/2022/Syksy/2022-09-16_EA_fi/attachments/index.html#10.A
   - » https://yle.fi/plus/abitreenit/2022/Syksy/2022-09-16_EA_fi/index.html#10

2) *Climate Crisis* text and item: https://yle.fi/plus/abitreenit/2022/Syksy/2022-09-16_EA_fi/index.html#12

APPENDIX 2.

ME censors', teachers', and learners' perceived challenges with ME items.

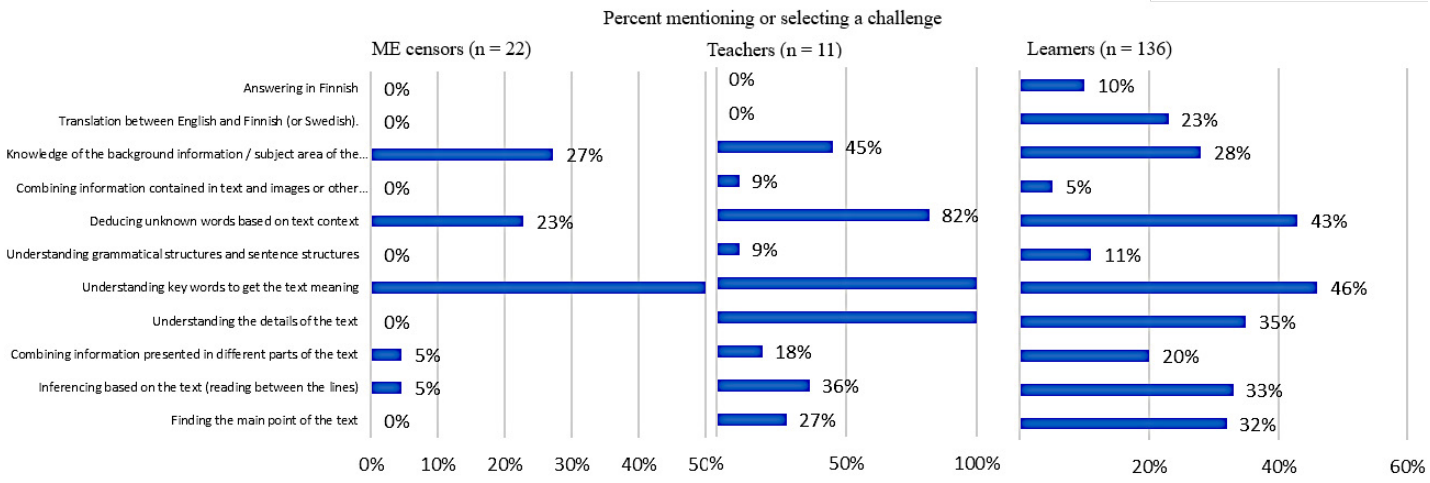

FIGURE 1. Percent of ME censors, teachers, and learners indicating particular challenges in the item *What is mentioned as a difference between GM and Tesla?*

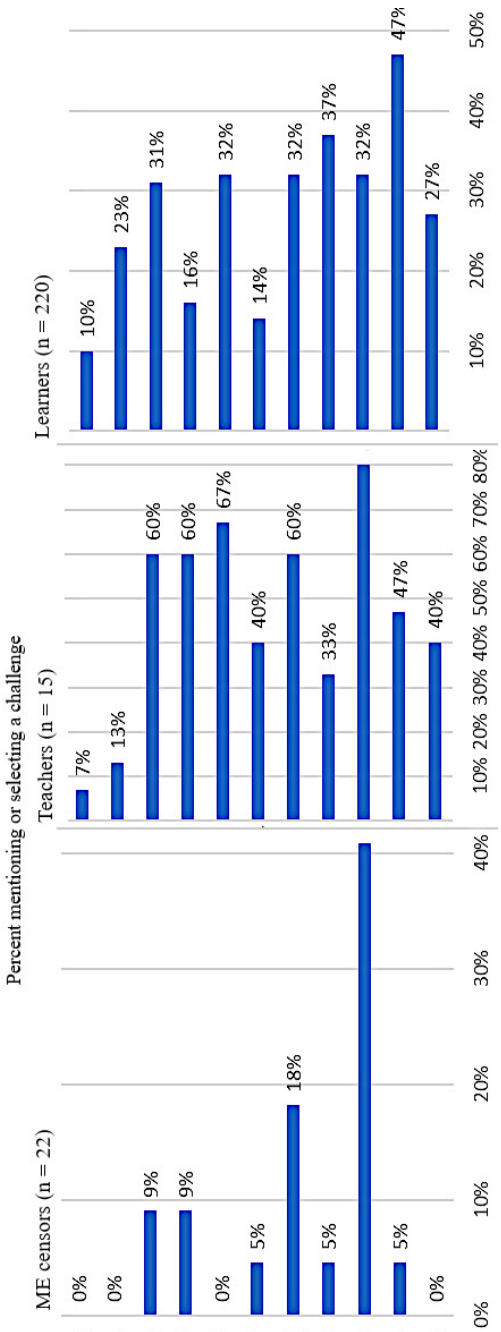

FIGURE 2. Percent of ME censors, teachers, and learners indicating particular challenges in the text *Climate Crisis.*