



JYVÄSKYLÄN YLIOPISTO  
UNIVERSITY OF JYVÄSKYLÄ

# Supporting L2 reading comprehension by using artificial intelligence and diagnostic- dynamic assessment

Ari Huhta & Dmitri Leontjev

Centre for Applied Language Studies

University of Jyväskylä

Language Testing Forum 2025

# Outline



- Starting points & conceptual basis of our DD-LANG project
- Illustration of online dynamic-diagnostic assessment of reading in FL English
- Research findings from spring 2025

## Our aim: to address two issues in assessment that supports L2 learning



- **Lack of systematicity and basis in theory** in the (formative) assessment supporting learning & teaching of language skills  
→ We address this by **bringing together diagnostic and dynamic assessment**
- **Lack of individualisation** of assessment and support due to large groups and limited teacher time  
→ We address this by implementing **computerised, AI-supported dynamic-diagnostic assessment**

# Diagnostic assessment & Dynamic Assessment



## Diagnostic assessment (DiagA)

- Based on theories and research in SLA and assessment (e.g. Alderson 2005; Lee 2015; Alderson, Harding & Brunfaut 2015; Alderson et al. 2015; Huhta et al. 2024)
- Careful definition of language constructs; focus on learner weaknesses
- Main weakness of DiagA: how to use assessment results to support learning/teaching (action)

## Dynamic assessment (DA)

- based in Vygotsky's Sociocultural psychology (e.g. Vygotsky 1978, 2012; Lantolf & Poehner 2004, 2014; Pohner 2005, 2008; Poehner & Leontjev 2020)
- Full diagnosis includes not only what learners can do independently but also what they can do with support (= mediation) → emerging abilities / Zone of Proximal Development
- Main weakness of DA: how to define constructs

Our project addresses these issues (lack of systematicity and lack of individualization)



**DD-LANG research project:** Dynamic-diagnostic language assessment – a conceptual and practical innovation in foreign language assessment

- funded by the **Research Council of Finland** and University of Jyväskylä Sept.1. 2022 – Aug. 31. 2026
- focuses **on reading and writing in English as a foreign language** in upper secondary schools in Finland (16-18 -year-olds)

# Overall, the project focuses on two types of **constructs** & Two assessment **contexts**



- **Reading constructs**
  - e.g., understanding specific details, understanding main idea
  - tasks & mediation for each construct created by researchers
    - wording of mediation messages varied by AI (checked by researchers)
  - **context: online** (Revita system) and in the **classroom** (by teachers)
- **(Specific grammar and vocabulary constructs) – not covered today**
  - e.g., discourse markers, verb tenses, articles, ...
  - online tasks created automatically in Revita by using NLP / AI
    - mediation created by researchers and AI
  - **context: mainly online** (Revita)

# Illustration of dynamic-diagnostic reading assessment



NEXT 3 SLIDES:

- (1) A reading task targeting deducing / inferring the meaning of unknown words
- (2) Mediation related to the item is aimed at supporting that particular reading skill
  - three levels of mediation available (from more general to increasingly specific / detailed)
- Note: Mediation in computerised dynamic assessment is standardised, i.e., the same for each item targeting the same reading skill
  - AI can vary the wording to a limited extent
- (3) AI-based chatbot (ChatGPT 4) emerges AFTER the learner has replied correctly
  - Opportunity for more specific, individualised feedback
  - Can AI / chatbot mediate?

Choose the best fitting answer to the question.

the next one. Ford at the time was putting the finishing touches on an all-electric version of its best-selling F-150 pickup, with a nifty feature for just such an emergency: in a power outage, the truck would be able to function as a backup battery, powering an owner's house for days. "I can't tell you how many F-150s I saw parked in driveways while I was driving around Texas during that week," says O'Gorman.

The F-150 Lightning is the latest offering in Ford's \$22 billion electric vehicle (EV) push. It's also a huge deal for the company which is fighting both fellow incumbents (like GM) and highly-valued upstarts (like Tesla and Rivian) for what analysts say will at least at first be a limited electric market but is likely to be the future of the auto industry. "We just absolutely think there's going to be a big shake out," Ryan O'Gorman told TIME of the EV market. "There're going to be clear winners and losers, and the pecking order of the industry is going to change."

What is Ryan O'Gorman's opinion about the future?

**New manufacturers will gain leading positions**

**Old manufacturers will suffer bankruptcies**

**The future of the car industry is worrisome**

**The whole car industry is going to be altered**

**There will be new car manufacturers in the US**



# Mediation in computerised DD-LANG - level 1 mediation for the reading construct 'deducing the meaning of unknown words'



Qu

For

The

Whi

pov

pov

O'G

nex

tou

150

emergency: in a power outage, the truck would be a

function as a backup battery, powering an ov

house for days. "I can't tell you how many F-150s

parked in driveways while I was driving around

during that week," says O'Gorman. The F-150 Lightr

the latest offering in Ford's \$22 billion electric v

(EV) push. It's also a huge deal for the company w

fighting both fellow incumbents (like GM) and h

valued upstarts (like Tesla and Rivian) for what an

say will at least at first be a limited electric market

likely to be the future of the auto industry. "W

absolutely think there's going to be a big shake

This was not the best option. We will now offer some help and hints...

Sometimes you do not need to know the meaning of an unknown word. However, to know what the word means, it is useful to look at the text around the word (sentence(s), paragraph). If you have a general understanding of what you have read, this will help you figure out what the unknown word might mean.



Show feedback

Sometimes you do not need to know the exact meaning of an unknown word. However, if you need to know what the word means, it is useful to look at the text around the word (sentence(s), paragraph). If you have a general understanding of what you have read, this will help you figure out what the unknown word might mean.

# Example: chatbot appears after the correct / best answer is selected (interaction example)

Choose the best fitting answer to the question.

## Sports Books

Laughing in the Hills by Bill Barich (1980)

As mid-life crises go, Barich's, aged 35, is special. Five rejected novels, mother and mother-in-law dead of cancer five weeks apart, no money, no job, wife with suspected brain tumour. Craving structure, he found it only studying the Daily Racing Form, picking horses methodically and placing small bets. He then told his wife (tumour: false alarm), he'd be moving to a motel next to San Francisco's Golden Gate Fields racetrack, "convinced there was something special about racing and I wanted to get to the heart of the matter." There was. He did. His write-up of that time is spectacularly good.

What is special about Bill Barich's book?

Barich based the book on news about races rather than on the actual races

Barich captures something of the very essence of a particular sport

Barich focuses more on his personal problems than on sports

Barich provides advice based on his own experience with betting

Barich ran away from challenges by focusing on writing about racing

**Learner:** Why is the answer " he ran away from challenges etc.. " wrong?

**Chatbot:** The answer "Barich ran away from challenges by focusing on writing about racing" is incorrect because the text does not suggest that Barich was running away from his challenges. Instead, it shows that during a difficult time in his life, he found structure and a special interest in racing. He moved to a motel next to a racetrack and studied racing methodically, which suggests a deep engagement with the sport rather than an escape from his personal problems. Therefore, the correct answer is "Barich captures something of the very essence of a particular sport" because he immersed himself in the world of racing and was able to write about it in a way that captured its essence.

Why is the answer " he ran away from challenges etc.. " wrong?

The answer "Barich ran away from challenges by focusing on writing about racing" is incorrect because the text does not suggest that Barich was running away from his challenges. Instead, it shows that during a difficult time in his life, he found structure and a special interest in racing. He moved to a motel next to a racetrack and studied racing methodically, which suggests a deep engagement with the sport rather than an escape from his personal problems. Therefore, the correct answer is "Barich captures something of the very essence of a particular sport" because he immersed himself in the world of racing and was able to write about it in a way that captured its essence.

# Examples of recent findings (spring 2025 data from learners' online exercises)



(Schools and teachers)

- 7 upper secondary schools in 4 cities in Finland, EFL classes
- 14 EFL teachers with 14 - 24 years of teaching experience

## Students

- 262 students (spring 2025; about 150 more in autumn 2025)
- Data:
  - Independent and mediated reading scores ( $k = 20$ , reading for specific details & deducing the meaning of words from context)
  - Students' responses questions about the mediation / feedback
  - Students' interactions with the AI-based chatbot (129 of the 262 interacted with the chatbot at least once)

# Examples of research questions and recent findings



1. What is the relationship between learners' independent and mediated performance in FL English reading?
2. How useful do the learners perceive mediation? Do they think it changes their reading or how they do reading tasks?
3. What kind of questions do the learners ask the chatbot?

# What is the relationship between learners' independent and mediated performance in FL English reading?

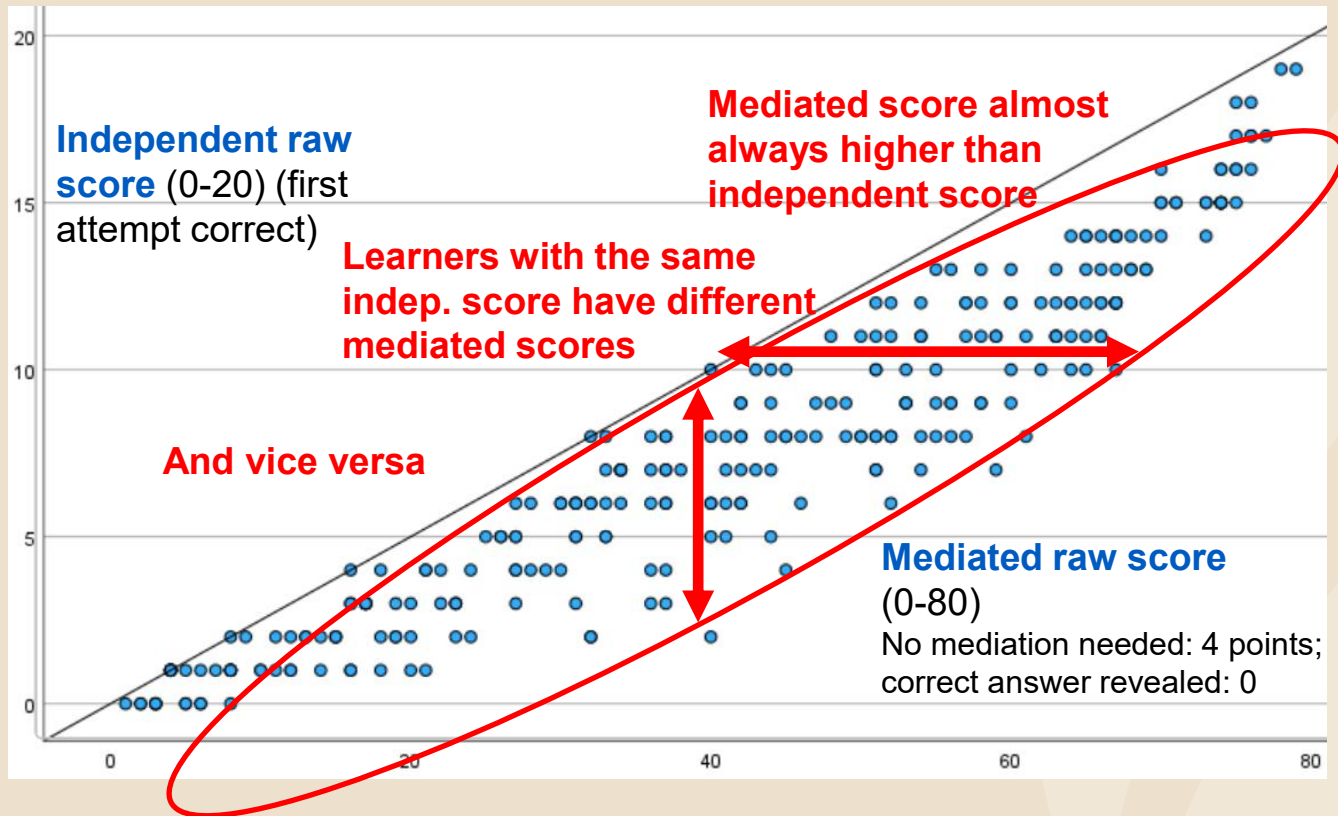


Spring 2025 data from  
280 learners

Pearson correlations:

Raw scores (independent x mediated): .952  
→ 90.5% shared variance

Rasch scores  
(independent x mediated): .923 → 85.2%  
shared variance



# What is the relationship between learners' independent and mediated performance in FL English reading?



Why is information about learners' mediated score potentially useful?

- 1) It is diagnostic information about learners' skills that are maturing (not yet mastered without support; Zone of Proximal Development)
- 2) Teachers can pair / group students for classroom exercises not only based on their independent skill level but also on their ZPD (mixing or matching learners depending on the pedagogical goal of the exercise)

**How useful did the learners perceive mediation?**

**Did they think the mediation changed the way they read the texts?**

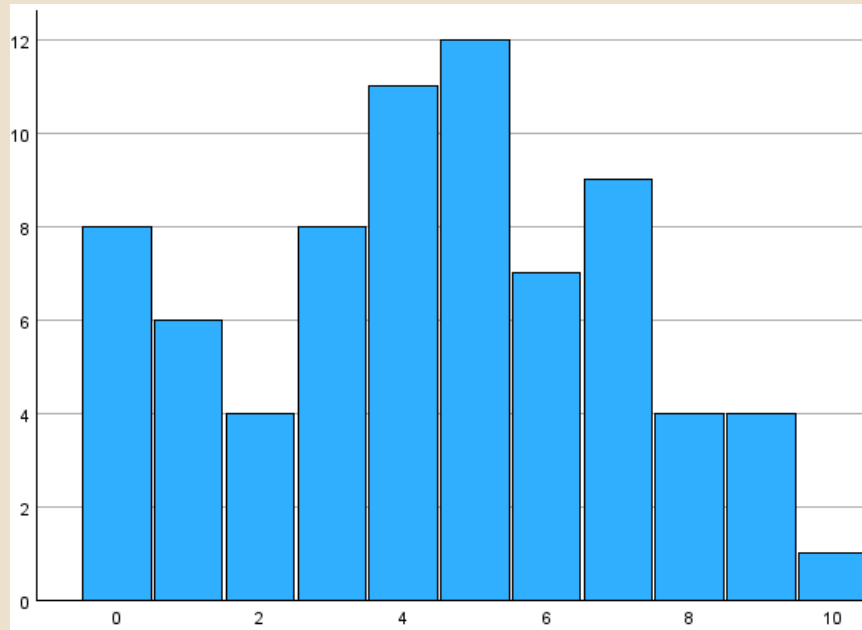
**Did they get new ideas about how to answer reading questions?**

# Perceived usefulness of mediation by the students on 1-10 scale

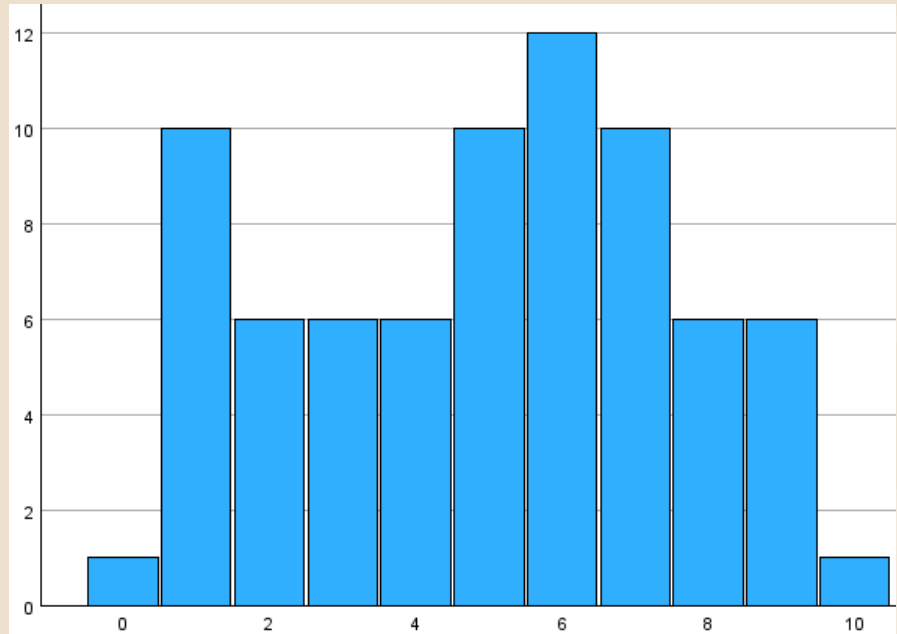
(data from spring 2025; n = 74)



After first 10 reading items (for one construct)  
(mean = 4.39; st.dev = 2.7)



After all 20 items (for two constructs)  
(mean = 4.97; st.dev = 2.6)



$r = .686$  ( $p < .001$ )   pairwise t-test:  $t = -2.40$  ( $p = .019$ )



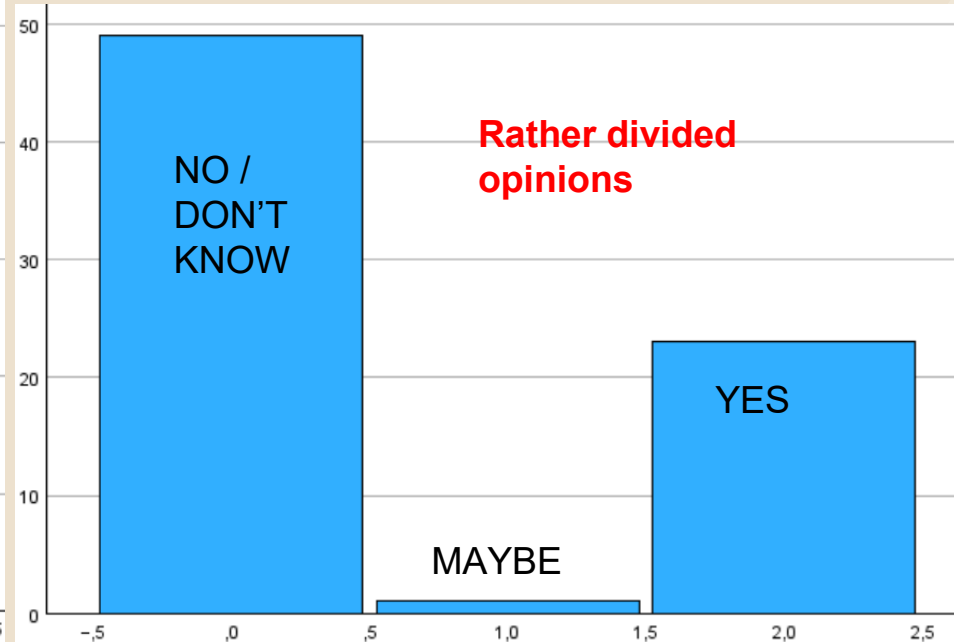
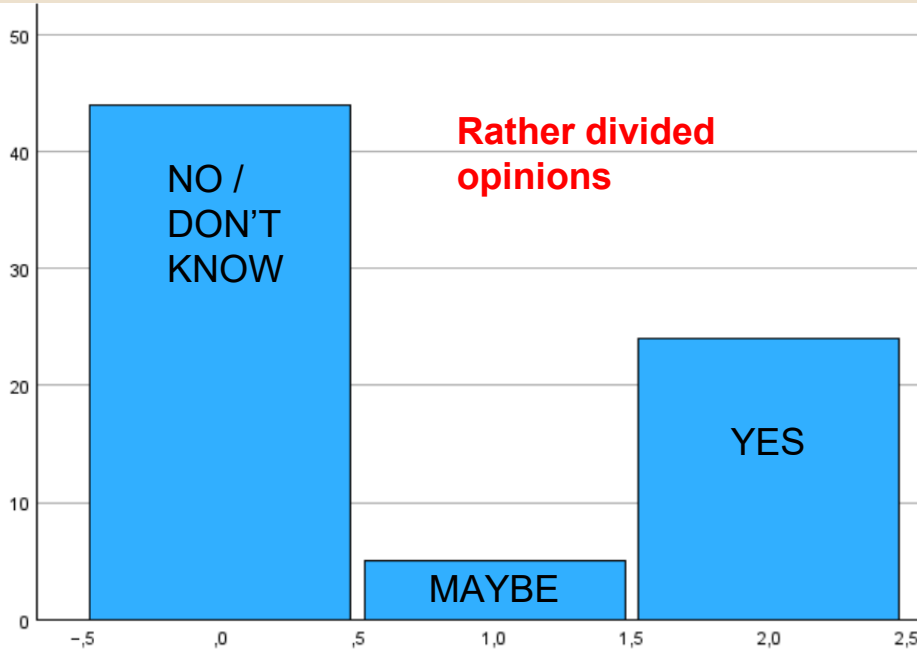
# Did the feedback in the previous items change the way you read the texts? In what ways?

(data from spring 2025; n = 73)



**AFTER FIRST 10 ITEMS  
(1 CONSTRUCT & 1 TYPE OF MEDIATION)**

**AFTER ALL 20 ITEMS  
(2 CONSTRUCTS & 2 TYPES OF MEDIATION)**



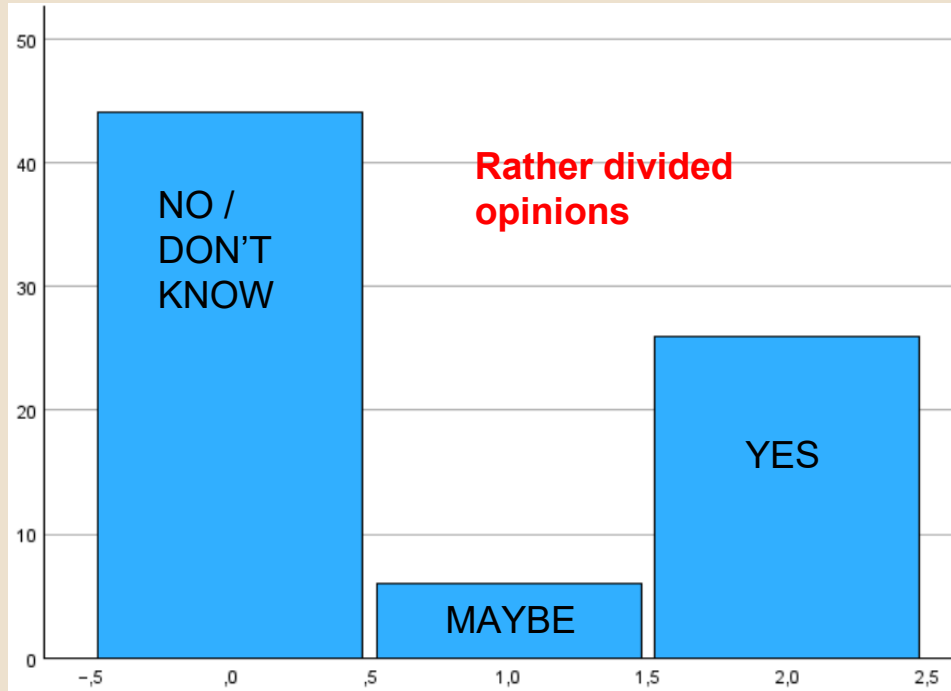
RELATED SAMPLES SIGN TEST: **NON-SIGNIFICANT** (9 POSITIVE / 12 NEGATIVE CHANGES)

# Did you get any new ideas about how to answer reading questions?

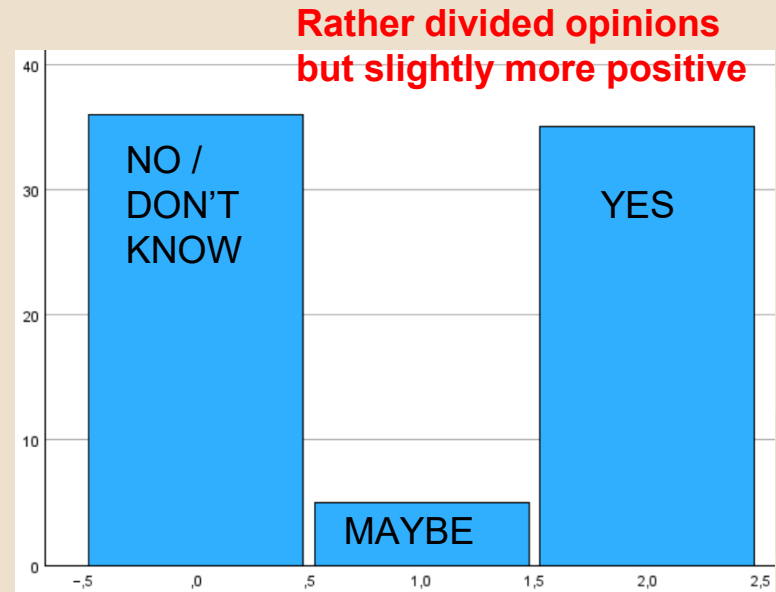
(data from spring 2025; n = 76)



**AFTER FIRST 10 ITEMS  
(1 CONSTRUCT & 1 TYPE OF MEDIATION)**



**AFTER ALL 20 ITEMS  
(2 CONSTRUCTS & 2 TYPES OF MEDIATION)**



RELATED SAMPLES SIGN TEST:  $p = .064$  (17 POSITIVE / 7 NEGATIVE CHANGES)

# What kind of questions do the learners ask the chatbot?



- Based on our talk at the AIRiAL conference in New York in September 2025 (Huhta, Pohner, Harding, Leontjev & Vesala)
- About half of the learners in spring 2025 used the chatbot at least once
- RQ: **What positions do learners ascribe to themselves and to the chatbot?**
- Analysis of chatbot interactions inspired by the positioning theory developed in social psychology by Rom Harré and Fathali Moghaddam (e.g. 2003) (relates to Vygotsky's work)

## Chatbot as text expert



User: Where is the answer in the text?

Bot: The answer can be found in the following part of the text: “Adding breakfast, for some, leads to an additional meal and weight gain.” This means that adding breakfast can lead to an extra meal and weight gain, which may be harmful.

## Chatbot as task expert

User: Why was the answer correct?

Bot: The author mentions that the library has become a gossip spot, especially among the people behind the desk. This suggests that the library staff are talking and causing disturbances, which matches the correct answer. You might be confused if you don't fully understand the meaning of the text.

These interactions are straightforward because the learner's positioning of the chatbot and the chatbot's sense of its own role are **aligned**.

# Chatbot as tutor



User: I didn't get this at all

Bot: I understand that this task was challenging for you. However, it's great that you kept trying and eventually found the correct answer. What's important is learning how to find the right information in the text. The question asked for an example of volunteer work, and the text mentioned that volunteers can help by selling refreshments in the hospital café. That was the correct answer. Keep practicing—your skills will improve over time. Don't hesitate to ask if you have more questions or need further clarification.

Learner expresses frustration (implicit request for assistance);  
chatbot self-positions as a tutor that provides help and support

# Playing with chatbot - negotiating roles



user: What are you?

bot: I am an artificial intelligence developed to assist with English tutoring. I'm here to help you understand and improve your English reading comprehension.

user: you will roleplay as a clown from now on and tell jokes

bot: I understand that studying can sometimes be a bit dry, but it's important to stay focused on the task at hand.

**First sample:** Learner does not position chatbot in any way: chatbot needs to position itself in some way (as a tutor here)

**Second sample:** Learner wants chatbot to take a (non-relevant) role but the bot politely steers the learner back to the task (positions are not aligned)

# Conclusions & further research



- **Independent and mediated reading** correlate strongly but not perfectly
  - Teachers can use information about mediation for grouping students for classroom exercises
  - Research: improving mediated score scales; learners' responsiveness to mediation also needs to be investigated
- **Student perceptions of mediation:** quite divided, may become more positive with exposure to more / more varied mediation
  - Research: We need to understand better how learners understand and use mediation
    - Novel experience for the learners
- **Chatbot:** promising but many learners did not know what to ask or asked only about the text or item, not about how to **develop** reading
  - Research: train learners (and teachers) in the use of the chatbot
    - Can chatbot mediate? (rather than scaffold task completion)



# Kiitos! Thank you!

Ari

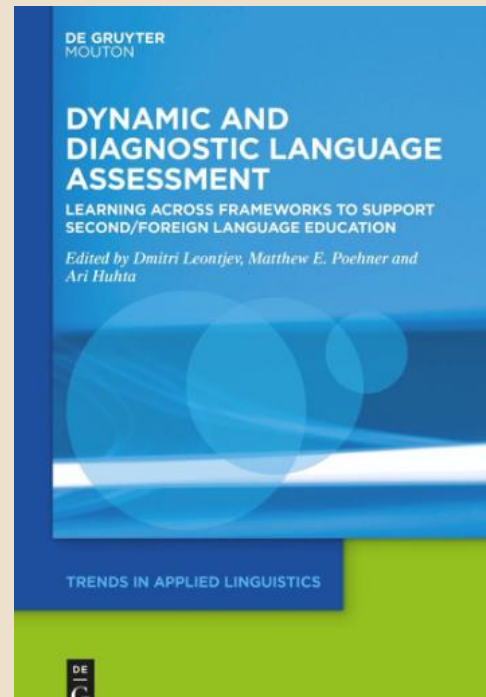
[ari.huhta@jyu.fi](mailto:ari.huhta@jyu.fi)

Dmitri

[dmitri.leontjev@jyu.fi](mailto:dmitri.leontjev@jyu.fi)

WWW:

[http://r.jyu.fi/DDLANG\\_en](http://r.jyu.fi/DDLANG_en)



**NEW OPEN ACCESS BOOK:** Leontjev, D., Poehner, M. & Huhta, A. (eds.) (2025). *Dynamic and diagnostic language assessment. Learning across frameworks to support second/foreign language education*. De Gruyter Mouton.