



Current trends in language assessment supporting learning and teaching: Combining the diagnostic and dynamic assessment frameworks

Ari Huhta

Centre for Applied Language Studies

University of Jyväskylä

ari.huhta@jyu.fi



Outline

- General trends in language testing and assessment over time
- Assessment(s) supporting learning and teaching
- DD-LANG project as an example of current developments in assessment supporting learning
 - Formative, diagnostic, cognitive diagnostic, and dynamic assessment
 - Their strengths and weaknesses
- Design, aims and stages of DD-LANG
- Examples of mediation in computerised reading tasks
- Findings so far
- Conclusions



General trends in language testing and assessment over time

- Most research on language assessment in the past focused on large-scale language examinations
- Focus was on the technical aspects of the examinations such as reliability but also on certain types of validity (e.g., predictive, content, internal/construct validity)
- Certain types of test were more typical than others (e.g., multiple choice, structured interview, essay/composition); statistical / quantitative analyses were typical (e.g. Davies 2014)
- Many things have changed over time that impact testing/assessment (e.g. Hamp-Lyons & Lynch 1998)
 - View of language (knowledge → use / communication)
 - View of validity & validation (validity types → holistic & new qualities)
 - Emphasis in assessment research (gatekeeping exams / assessment of learning → wider scope, e.g., assessment for learning)
- Today, I focus on the increasing importance of assessments that support L2 learning



Assessment(s) supporting learning and teaching

- Formative Assessment
 - Assessment for / as learning, learning-oriented assessment, ...
- Diagnostic Assessment (based on SLA & language testing research)
- Cognitive Diagnostic Assessment
- Dynamic Assessment



Road to the DD-LANG research project

Formative assessment; Assessment for/as learning; Learning oriented assm.

e.g. Scriven 1967; Bloom et al 1971; Black & Wiliam 1998; Turner & Purpura 2016; Purpura & Turner 2023
Education Language assessment Pedagogy

Diagnostic assessment *DIALANG DIALUKI-project*

e.g. Spolsky 1992; Shohamy 1992; Alderson 2005; Huhta 2008; Alderson et al 2015, 2016; Harding, Alderson & Brunfaut 2015; Alderson, Brunfaut & Harding 2015; Huhta et al 2024
LTRC main theme 2013 & Language Testing special issue 2015
SLA research Language testing research CEFR-related research (SLATE)

Cognitive Diagnostic Assessment

e.g. Buck et al 1997; Buck & Tatsuoka 1998; Jang 2009; Mei & Chen 2022
Educational measurement Cognitice psychology

Dynamic assessment

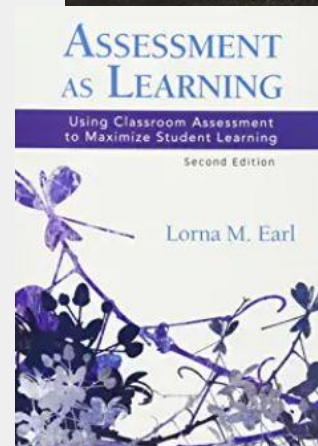
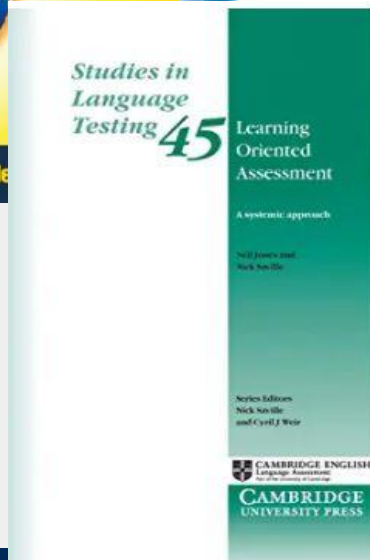
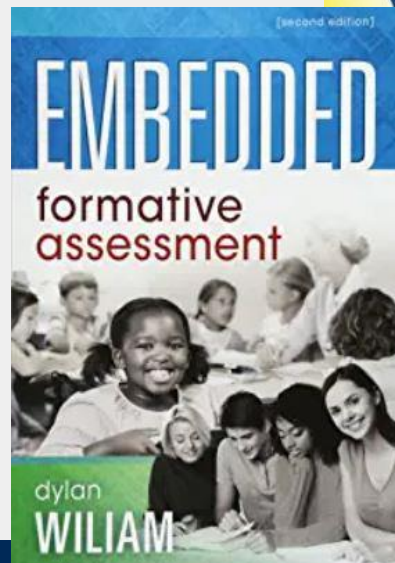
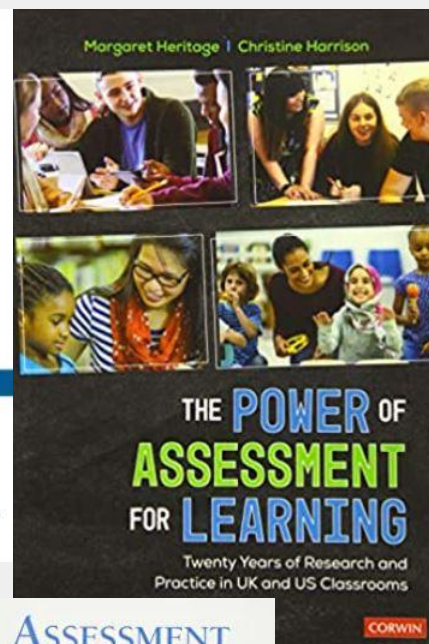
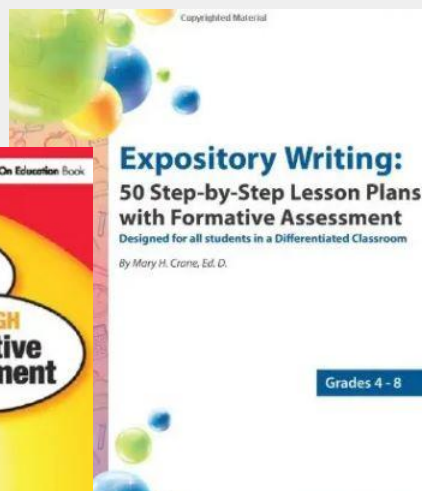
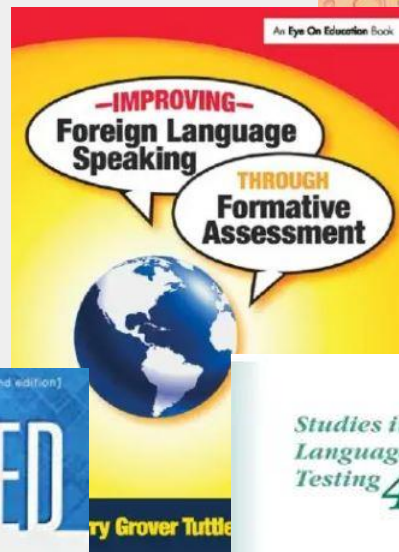
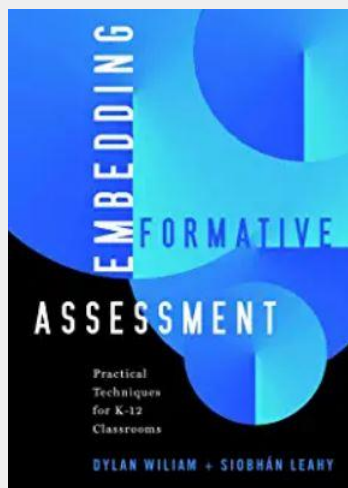
e.g. Vygotsky 1978; Lantolf & Poehner 2005; Poehner & Lantolf 2013; Poehner & Leontjev 2020
Sociocultural theory Developmental psychology

**DD-LANG
project**



Formative assessment

- A general term to refer to assessment that supports learning directly (coined by Scriven 1967) (e.g. Sadler 1989; Rea-Dickins & Gardner 2000; Leung & Mohan 2004; Wiliam 2011)
- Other closely related terms and approaches:
 - Assessment for learning / assessment as learning (e.g. Laveult & Allal 2016)
 - Learning-oriented assessment (LOA) (Purpura & Turner 2023)
- Typical features:
 - Typically in the classroom by the teacher (also self- and peer-assessment)
 - Multiple, varied methods (observation, quizzes, tests, portfolios, learning diaries)
 - Continuous
 - Provides feedback to learners
- If done systematically, can improve learning significantly (e.g. Black & Wiliam 1998)





Formative assessment

- strengths and weaknesses

- Strengths
 - Can be implemented in many teaching/learning contexts
 - When e.g. a teacher works with learners for longer periods of time, they get to know the learners well
 - When done systematically by using multiple approaches and over time, FA covers the skills of interest better than any single test can (construct coverage better)
- Weaknesses
 - Usually not informed by theory and research on the skills of interest (and how they develop); the curriculum / course plan / textbook is the only/main guide
 - Difficult with large groups and/or with unfamiliar learners
 - Difficult to gather reliable, systematic information about all learners by observing learners in the classroom



Diagnostic Language Assessment (SLA & language testing based)

- Started with the design of the DIALANG system (1997-2004) (Alderson 2005), multilingual online system for diagnosis and feedback (14 languages, 5 skills); language tests, self-assessment and rich feedback and advice for improvement
- Led to research into the use of SLA and language testing research for diagnosis and to the design of other diagnostic tests

Characteristics of diagnostic testing: aims to identify learner strengths & weaknesses (Alderson 2005); clear definition of constructs to be measured; and detailed feedback (Jang & Wagner 2014)

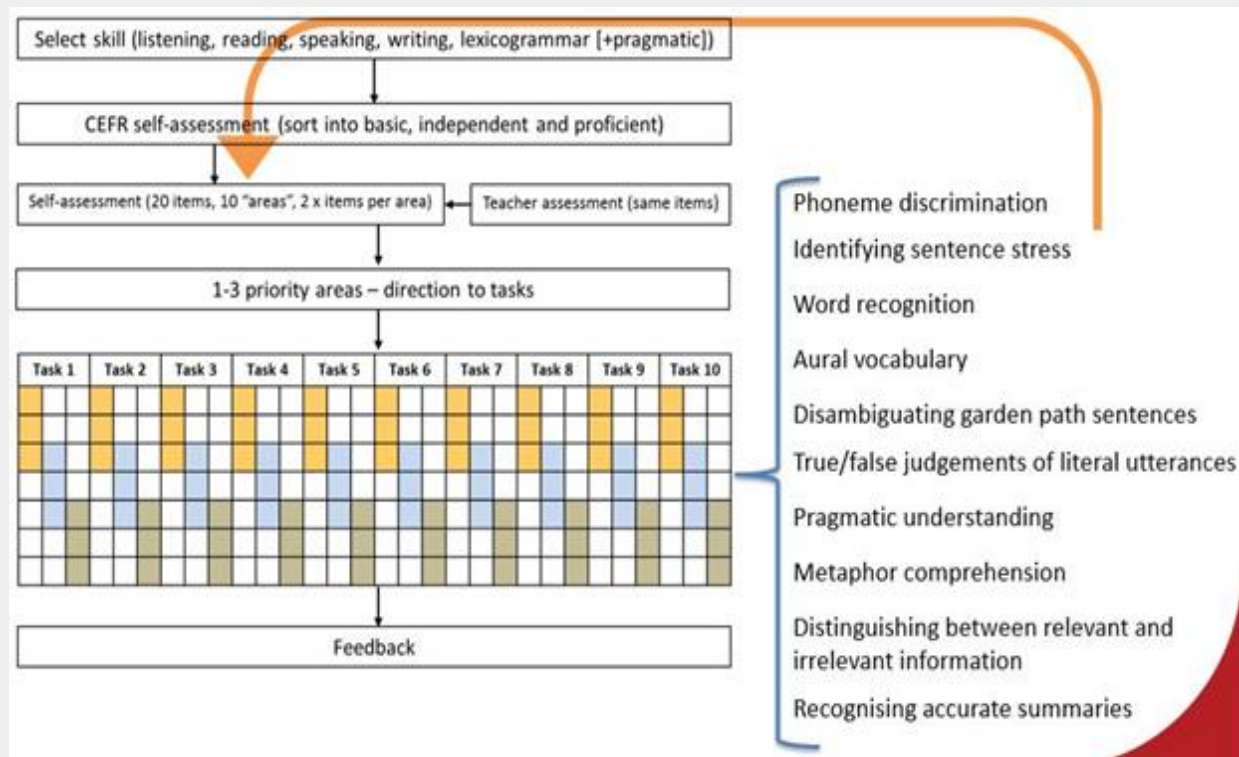
L2 Diagnostic assessment has been further elaborated conceptually and through research since 2000s, particularly by Alderson & colleagues at Lancaster U. and U. of Jyväskylä (e.g., Alderson, Haapakangas et al. 2015; Harding, Alderson, & Brunfaut 2015; Harding et al. 2018; Huhta 2008; 2023; Huhta, Harsch, Leontjev & Nieminen 2024)



Phases of diagnosis: 'Idealized Diagnostic Procedure'

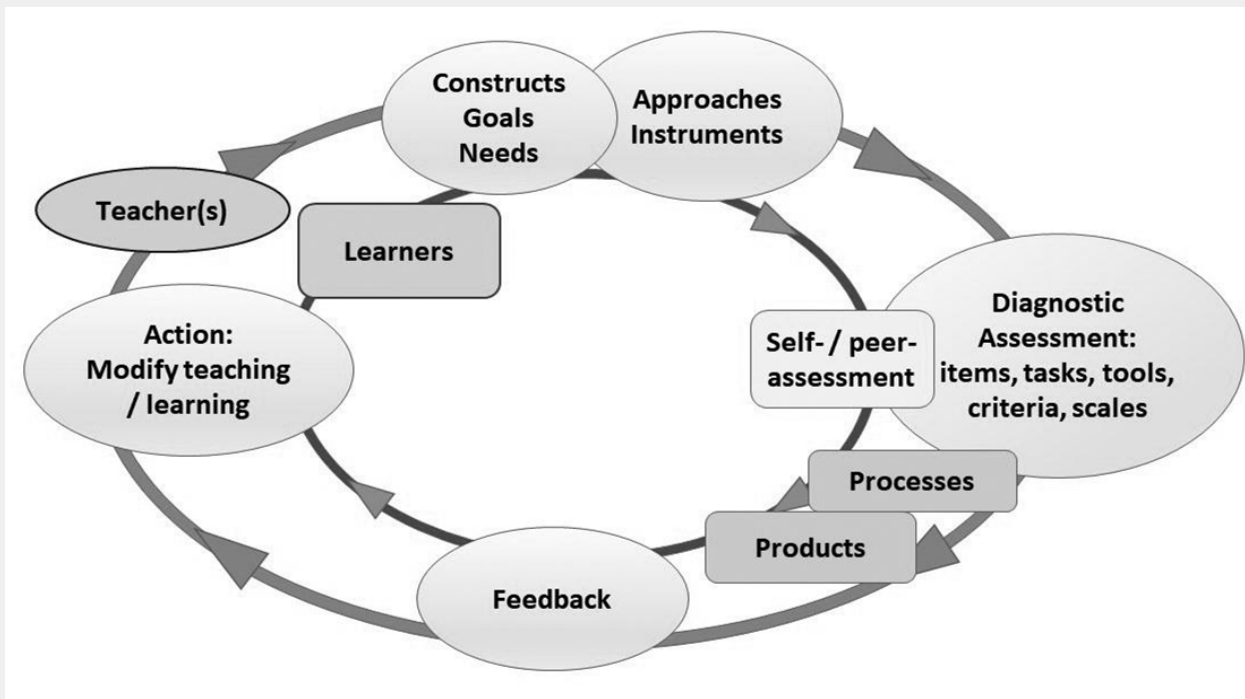
(Harding, Alderson, & Brunfaut, 2015):

- 1) Listening & observing
- 2) Initial assessment (hypothesis formulation)
- 3) Hypothesis verification
- 4) Diagnostic decision & feedback



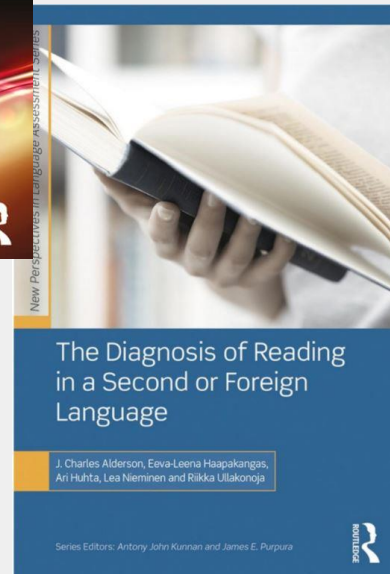
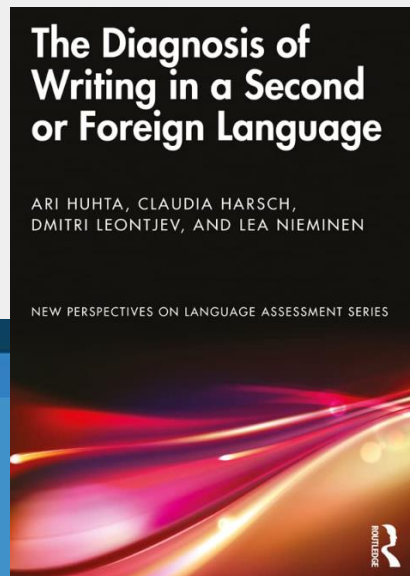
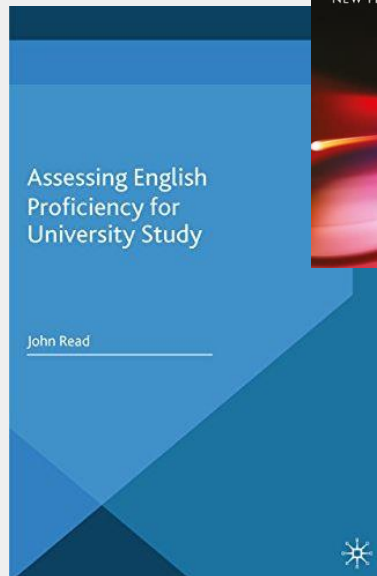
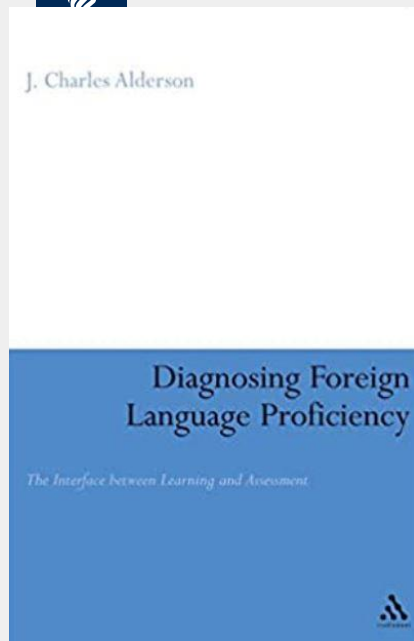


Diagnostic cycle



Huhta, Harsch, Leontjev, Nieminen (2024)

Books on diagnosing reading and writing





Diagnostic assessment (SLA based)

- strengths and weaknesses

- Strengths
 - (More) informed by theory and research on (1) assessment and (2) the skills of interest (and how they develop)
 - Carefully designed stages and procedures (e.g. tests) are sometimes available
 - Technology makes diagnosis more practical also with large groups and with unfamiliar learners
- Weaknesses
 - Does not automatically inform teachers/learners about how to use assessment results in teaching/learning (low impact)
 - Still relatively few diagnostic tests available for L2 skills

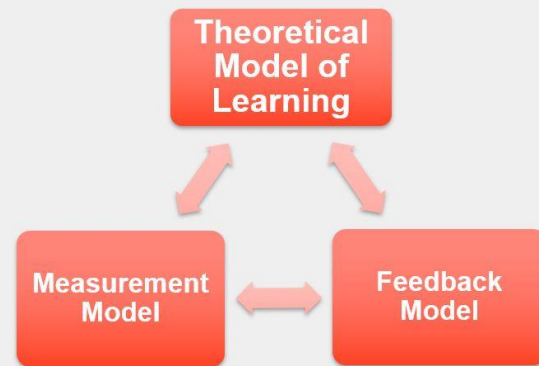


Cognitive Diagnostic Assessment

Cognitive Diagnostic Assessment (CDA) is assessment that

- was developed in cognitive psychology and educational measurement
- determines learners' strengths and areas of improvement;
- provides diagnostic feedback for learners

(for L2, see e.g. Jang 2009; Mei & Chen 2022)



Stages

1. Definition of the constructs to be learned → **theoretical model of learning**
2. Design & analysis of test items in terms of the detailed definitions → **measurement model**
 - In CDA, one item can measure more than one (sub)construct
3. Administration of the CDA test
4. Statistical analyses of the test data (multidimensional latent class models)
5. Reporting detailed results to the learner, teacher, ... → **feedback model**



	Comprehending explicit textual info	Processing syntactic knowledge	Processing vocabulary knowledge
Item1	1		
Item1	1	1	
...	1		1
Item10			
Item11		1	1
...		1	1
Item20	1		



Diagnostic scoring & reporting / feedback from CDA tests

	Comprehending explicit textual information	Processing syntactic knowledge	Processing vocabulary knowledge
Learner 1	70%	50%	70%
Learner 2	50%	80%	40%
Learner 3	90%	80%	80%



Cognitive Diagnostic Assessment

- strengths and weaknesses

- Strengths (similar to the SLA based diagnostic assessment)
 - (More) informed by theory and research on (1) assessment and (2) the skills of interest (and how they develop)
 - Carefully designed stages and procedures (e.g. tests) are sometimes available
 - Technology makes diagnosis more practical also with large groups and with unfamiliar learners
- Weaknesses (partly the same as in SLA based diagnostic assessment)
 - Does not automatically inform teachers/learners about how to use assessment results in teaching/learning (low impact)
 - Only a few diagnostic tests available for L2 skills
 - Developing CDA requires considerable expertise in statistical analyses



Dynamic Assessment (DA)

- Since the 1950s / 1960s, Dynamic Assessment (DA) has been used as a diagnostic tool primarily in psychological research and special education and then in educational research on teaching / learning of school subjects
- Dynamic Assessment has begun to be applied in L2 research and practice since the 2000s (Poehner 2008; Lantolf & Poehner 2004; 2013)



Dynamic Assessment

- DA is based on a sociocultural approach to learning
- Vygotsky's theory of the Zone of Proximal Development (ZPD) is a key concept in DA
- ZPD is **the distance between the actual developmental level** as determined by independent problem-solving and **the level of potential [proximal] development** as determined by problem-solving under adult guidance or in collaboration with more capable peers (Vygotsky, 1978: 86)
- Learning happens in the zone of proximal development



Forms of Dynamic Assessment

In general, DA integrates assessment into learning (and teaching)

- Learning happens (best) with **mediated** learning activities, i.e. in interaction with, and supported by, the teacher (or a more capable peer) – or by computer

Interactionist Dynamic Assessment

→ DA is mostly applied in classroom contexts where learners and teachers can interact face-to-face

- teachers can find out learners' strengths and problems through **mediation and collaboration** with learners
- at the same time, the teacher helps learners overcome these problems

Interventionist Dynamic Assessment

- Computer-based tests based on DA where feedback guides the learner through the tasks by becoming more explicit if the learner fails to respond correctly (Lantolf & Poehner 2013; Leontjev 2016)



How does Dynamic Assessment differ from traditional assessment?

Traditional assessment (including diagnostic assessment / testing)

- Test presents items and the learner is asked to respond to them successively, without feedback or intervention until the test is over
 - Feedback is given afterwards as e.g. a score or verbal comments
 - Traditional assessment measures learner's ability to perform **independently**

Dynamic assessment

- In interaction, the teacher (or computer) can first find out what the learner can do alone, independently and, then, what the learner can do with assistance (with mediation)
 - DA measures learner's **mediated** performance
- Prediction: What the learner can do now with assistance is what they will be able to do independently in the future



Dynamic Assessment

- strengths and weaknesses

- Strengths
 - Based on a clear theory of development (Sociocultural Theory of cognitive development)
 - Combines assessment with teaching (and learning) during assessment sessions
 - Expands the scope of assessment (diagnosis) to skills in the process of maturing (ZDP)
 - Particularly useful in face-to-face / one-to-one settings
 - Developments exist also for L2 and for computerised DA
- Weaknesses
 - In L2 contexts, often no clear view of the skills of interest (constructs)
 - Still relatively difficult to implement on a computer (for use with larger groups)
 - Mediation in computerised DA is quite inflexible



DD-LANG PROJECT



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

- (1) **Lack of systematicity** and **basis in theory** of learning / development in the skill(s) of interest
- (2) **Lack of individualisation** due to large groups and limited teacher time

→ potential issues for the quality and impact of assessment, including fairness

→ possible consequences include learners not receiving instruction that optimally supports their language development



What do we do to address these issues?



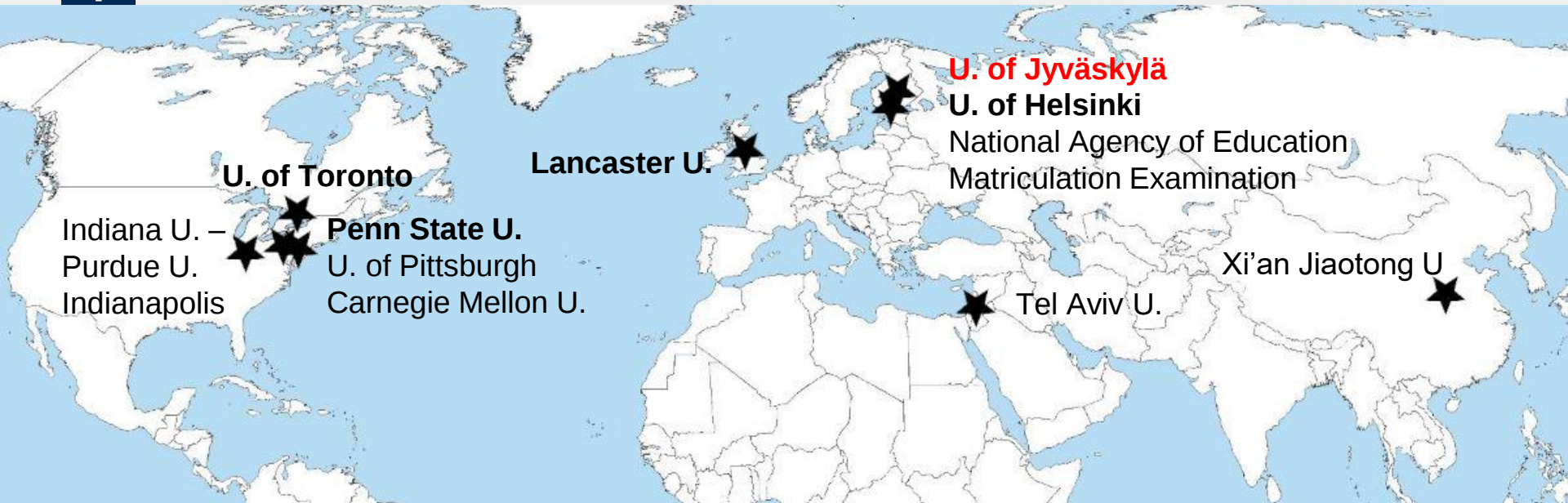
We conduct research that aims to advance the **theoretical, empirical and practical basis** of L2 assessment for learning

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)

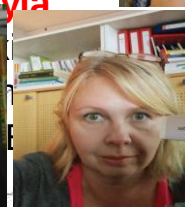


Collaborators





Collaborators



Eunice Jang (U.Toronto)
Matthew Poehner (Penn State)
Remi van Compernelle (Carnegie M.)
Marta Antón (Indiana U.)
Richard Donato (U. Pittsburgh)

Luke Harding & Tineke
Brunfaut (Lancaster U.)
Tziona Levi (Tel Aviv U. &
Min. of Education)
Yaru Meng (Jiaotong U.)

Ari Huhta, Dmitri Leontjev, Pirjo Pollari,
Askto Tolvanen (U. Jyväskylä)
Roman Yangarber, Raili Hildén (U. Helsinki)
Anu Halvari (National Agency)



How do we address these issues in DD-LANG?

(systematicity, theory-based, individualisation)

(1) By integrating L2 dynamic assessment (DA) and diagnostic assessment (DiagA)

→ clearer theoretical underpinnings of

(a) assessment **constructs** (& what instruction is intended to promote)
(contribution of DiagA)

(b) **development** of learner L2 abilities (contribution of DA)

DA: Integration of mediation, interpretation of learner responsiveness to expand diagnosis to include developed abilities + emerging/partly formed abilities/understandings



How do we address these issues in DD-LANG?

(2) By increasing synergy of teacher-led classroom assessment & AI-enhanced online assessment (that implements dynamic and diagnostic assessment)

→ online assessment provides teachers and learners with more, and more individualised, information about learner development

→ online assessment tasks provide input to classroom activities

Overall, DD-LANG aims to provide better opportunities for learner L2 development, fairness through meeting the needs of all learners, and, supporting all learners for success



DD-LANG covers:

Two types of constructs & Two assessment contexts

- **Broader (reading) constructs**
 - e.g., inferring word meanings from context, understanding main idea
 - tasks & mediation for each construct created by researchers
 - **context:** online and in the classroom
- **Narrower constructs**
 - e.g., discourse markers, verb tenses, articles, ...
 - online tasks created automatically by using NLP / AI
 - mediation created mostly by researchers
 - **context:** mainly online



Steps: (1) Defining (reading) constructs; (2) Designing mediation for each construct

Sources of information for defining reading constructs:

- Review of theories and research on L2 reading
- Online surveys of key stakeholders: item writers ($n=22$), teachers ($n=46$), students ($n=725$)
- Teacher interviews ($n=8$)
- Think-aloud protocols with expert readers ($n=6$)

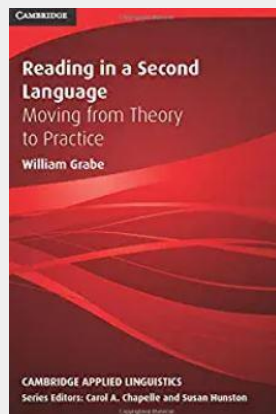
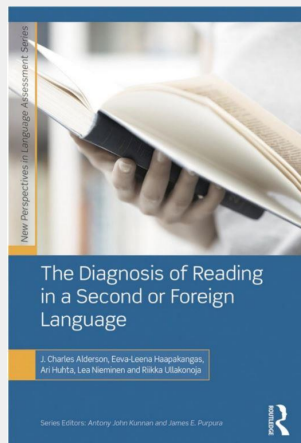
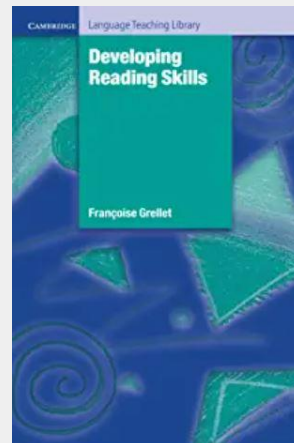
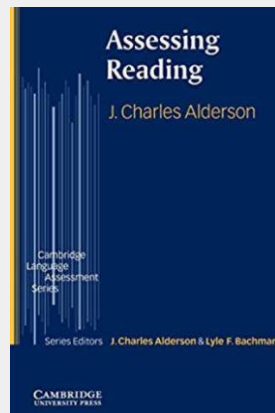
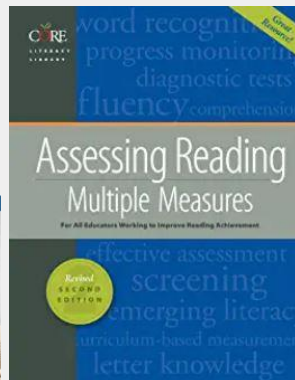
Designing mediation & One-on-one piloting of the mediation



Where to get ideas for assessing / diagnosing reading?

Reading in a Foreign Language

A refereed international online journal of issues in foreign language reading and literacy





Research questions of DD-LANG

The 5 research questions:

1. In what ways does Finnish Gymnasium students' **reading** and **writing** ability in English improve during the study?

→ **Impact on language proficiency**

2. How useful do teachers and learners find the diagnostic profiles of learner abilities that emerge from the computerized DD-LANG?

→ **Impact on e.g. how reading and writing are understood**

RQ1 & RQ2 → we need to define in enough detail what we mean by R & W and then operationalise these skills / constructs



Research questions (cont.)

3. Are teachers' assessment practices changed following their participation in training to use DD-LANG in the classroom? If so, how?

→ Impact on assessment practices

4. Does their experience with the DD-LANG framework change teachers' and learners' beliefs regarding language teaching, learning, and assessment?

→ Impact on beliefs about the key aspects of language education

5. Based on the findings, in what ways are DA and Diag-A each enriched through the proposed integrated framework? How viable is the integrated DD-LANG framework for use in other contexts?

→ Theoretical contributions



Design

- **Longitudinal intervention** (pre-post control group design)
- **QUAN**: measures of reading and writing; online activities; **QUAN/QUAL**: questionnaires about practices and beliefs; **QUAL**: interviews; classroom observation; training sessions and discussions with teachers (informed by Vygotskian **praxis**)
- Experimental group: **dynamic-diagnostic assessment** and go through the enrichment programme in the **Revita system** (the U. of Helsinki) and the classroom
- Teachers in the experimental group will be **trained in the use of dynamic assessment**
- Control group: Revita system but without mediation



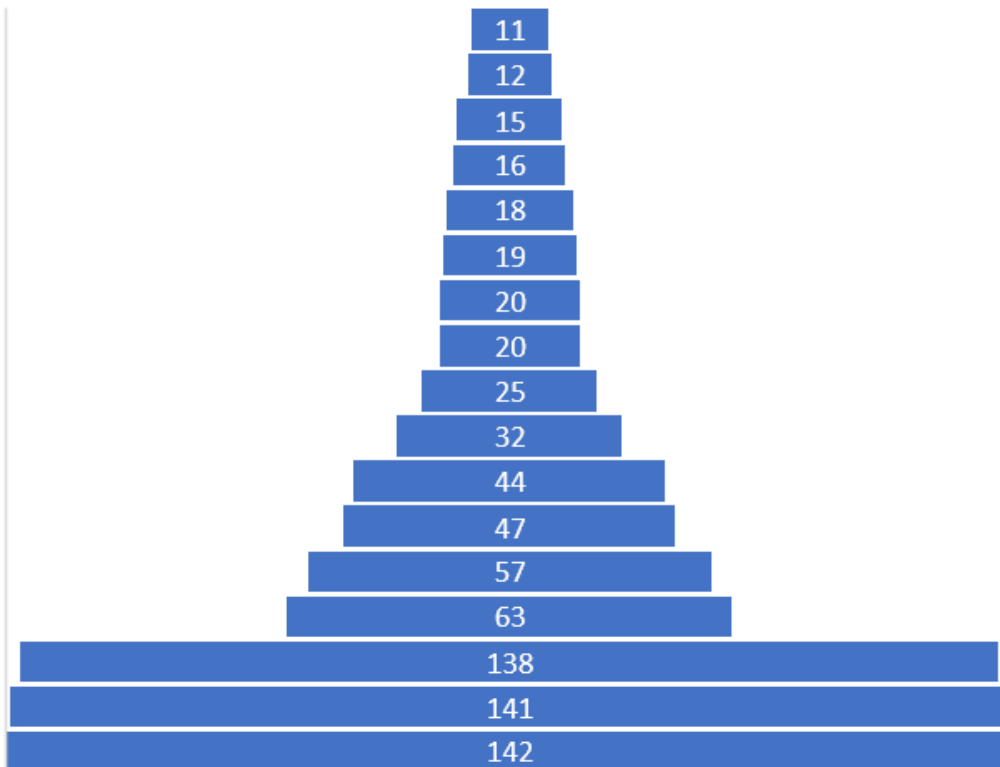
Some findings so far: Illustrative examples



Construct definitions: Online survey of students (n=725) “Think about a person who is a good reader in English. What can such a person do in your opinion?”

Ajattele henkilöä, joka on mielestäsi hyvä lukemaan englanniksi. Mitä tällainen henkilö osaa mielestäsi tehdä?

- nativelikeness
- ability to construct sentences
- speaking (more)
- reading books/novels
- proficient/experienced writer
- proficient/experienced speaker
- ability to find and use relevant...
- reading/writing a variety of texts
- knowledge of grammar, syntax and...
- understanding the main idea of the...
- communication skills
- reading (more)
- Accuracy
- Proficient/experienced reader
- reading comprehension
- pronunciation
- Vocabulary knowledge

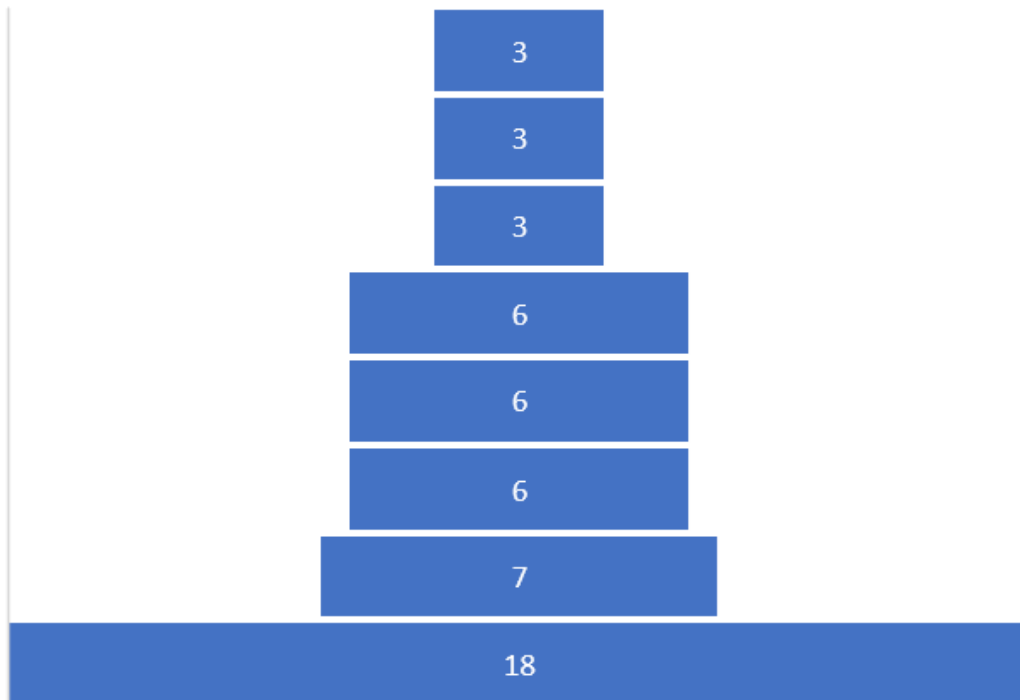




Construct definitions: Online survey of English language teachers (n=39)

What should a Lukio student be able to do/know to be a good reader in English?

- critical reading
- finding connections between ideas i...
- understanding the main idea of the...
- deducing meaning of words
- knowledge of genres, text types, styl...
- reading/writing a variety of texts
- knowledge of grammar, syntax and...
- Vocabulary knowledge



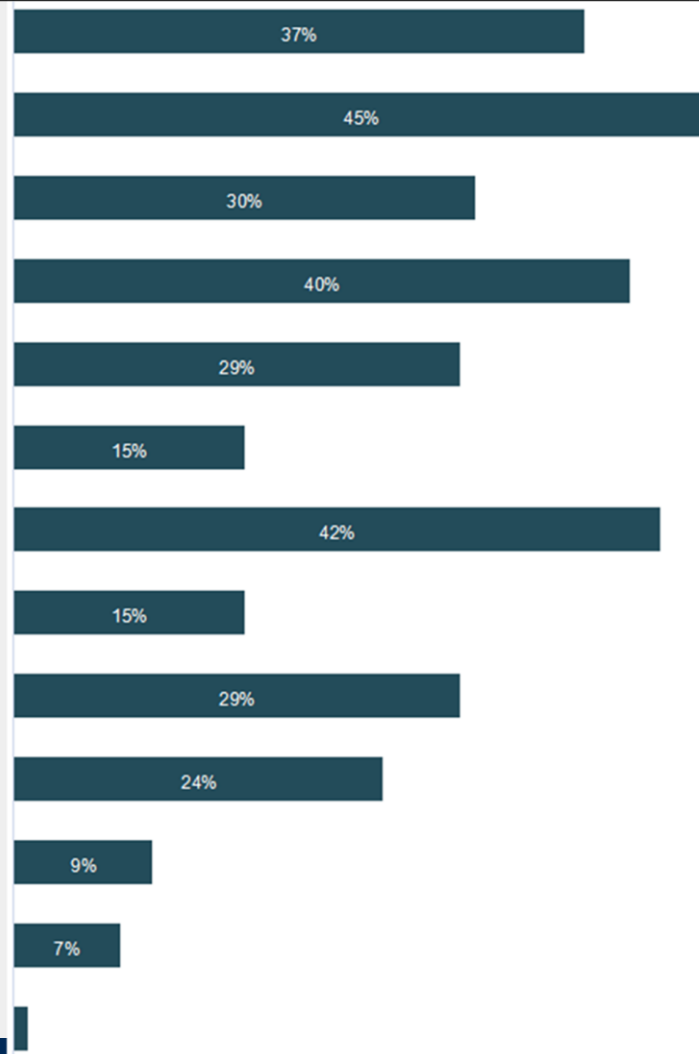


What challenges have you faced when answering this item?

What significant past experience does the narrator mention?

- Having a public post
- Teaching classes
- Sitting an exam

- Finding the main idea of the text
- Reading between the lines
- Connecting information from different text parts
- Understanding separate details in the text
- Understanding the meaning of key words
- Understanding grammatical and sentence structures
- Deducing the meaning of unknown words from the context
- Combining information in text with images or graphs
- Background knowledge
- Translation between English and Finnish (or Swedish).
- Answering in Finnish
- Other
- Other





Designing mediation for reading skills (constructs)



Sample Reading Item (inferring word meanings)

Electric Car

Ford's New All-Electric F-150 Lightning Is More Than It Seems

While driving around suburban Houston during Texas' power crisis in February 2021, which left millions without power for days, Ford energy services lead Ryan O'Gorman was thinking less about that disaster than 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

What is mentioned as a difference between GM and Tesla?

GM has a more established position

GM is thought to focus on petrol engines

GM keeps fighting fellow incumbents

GM will decide the future of the industry

GM's reputation is more disputed



Mediation in computerised DD-LANG - level 1 mediation

Qu
For
The
Whi
pow
pow
O'G
nex
tou
150

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

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.



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.



Mediation in computerised DD-LANG - level 2 mediation

The following may help you find the meaning of an unknown word:

- **Which word comes right before / after** the unknown word? Does that word help you understand the difficult word?
- Can you **divide the unknown word into parts**? Can you recognise the meaning of the parts? (e.g., un-help-ful)
- Recognising only part of the meaning may help, too. Is the word **positive or negative**?
- How are details in a sentence or in several sentences connected? For example, the word 'but' suggests that details are contrasted, and 'because' tells us that a reason is given. Look for these kinds of words in the text.

Tesla?

Show feedback

GM has a more established position

GM is thought to focus on petrol engines

GM keeps fighting fellow incumbents

GM will decide the future of the industry

GM's reputation is more disputed



Mediation in computerised DD-LANG - level 3 mediation (note: this level of mediation is item specific, not general / construct specific)

To find the answer to this question, you should think how GM and Tesla are compared to Ford:

- Think what kind of company Ford is.
- Words like "fellow" tell us about GM in relation to Ford.
- "both... and..." signals that GM is a company like Ford whereas Tesla is different.

Remember that the goal is not to understand the full meaning of all the words but only as much as needed to answer the question.

Show feedback

What is mentioned as a difference between GM and Tesla?

GM has a more established position

GM is thought to focus on petrol engines

GM keeps fighting fellow incumbents

GM will decide the future of the industry

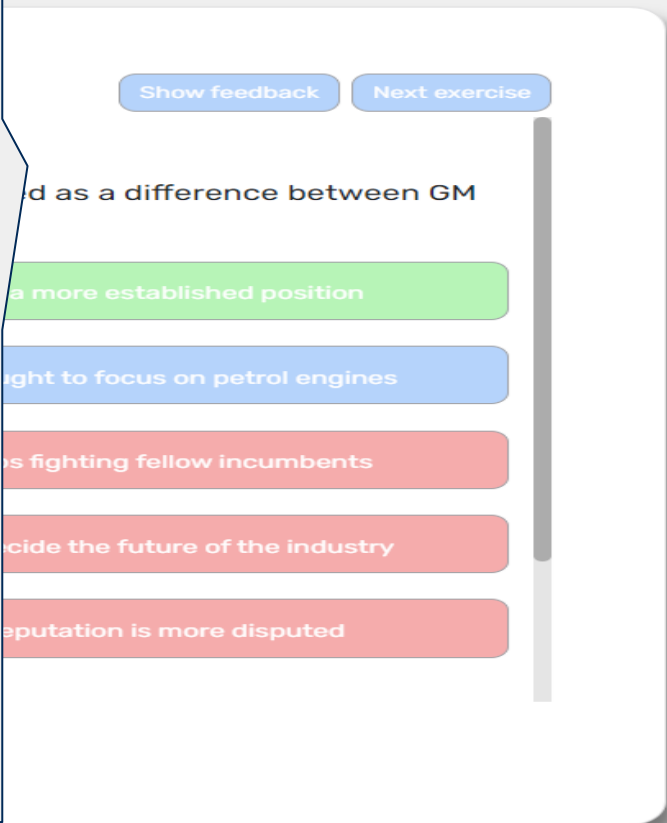
GM's reputation is more disputed



Mediation in computerised DD-LANG - final mediation (synthesis of what the learner can do to infer word meanings)

There are many things in the text that can help you figure out the meanings of unknown words in the future:

- **Start** by looking at the **title and headings, pictures, and keywords**. They can help you understand the general idea of the text.
- You can also try to **read the entire text**, paragraph, or sentence. A general understanding of the text can help you figure out the meaning of unknown words.
- Look at what **part of speech** it is, for example, a verb (to go) or an adjective (beautiful). That can help you understand its meaning.
- Look at the **text around the unknown words**. Pay attention to linking words that tell us how different ideas are connected. Check the words that are used together with the unknown words. They can give you an idea of what it means.





Creating automatic exercises for narrower constructs (grammar & vocabulary) with mediation in the Revita system

Steps:

- Learner (or teacher / researcher) selects a text and inputs it into Revita
- Learner (or teacher / researcher) selects the (narrow) construct that they want to train
- Revita uses NLP to create exercises from the text that focus on chosen construct (e.g., gap-fill tasks with or without multiple-choice options)
- Learner takes the exercises and when they make mistakes, Revita displays researcher-designed mediation messages until the exercises are completed
- Revita monitors learner progress across the chosen constructs, including amount of mediation (and changes in mediation)



Example of a narrow construct: Discourse markers

There was a feeling I got before I spoke in front of an audience and sometimes also before an event that was less public but still important, an event that could have consequences in my life – taking the LSATs, for example, I in a classroom on the campus of Harvard.

Check

therefore

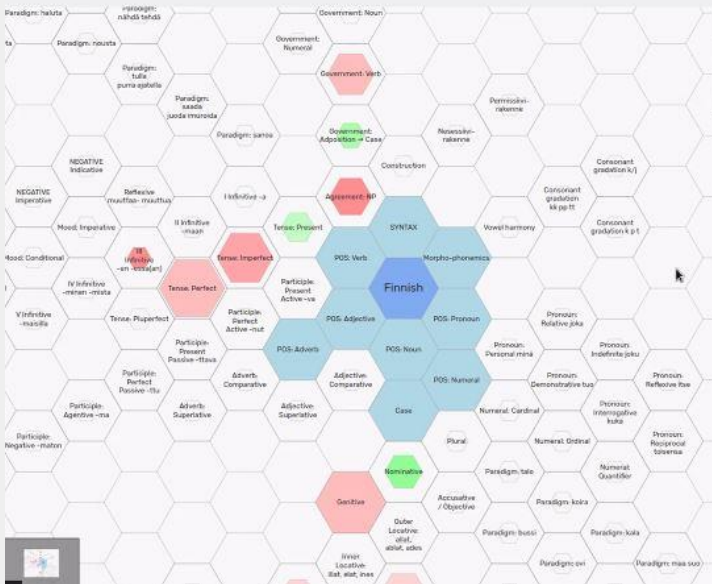
unlike

which

whom

Next snippet ↓

 Report a problem



Each hexagon =
topic / unit / skill
area

Green / blue = mastered
Yellow = partial mastery
Red = not mastered
White = no evidence yet



Conclusions and next steps

- Differing needs & abilities require different instruction. DD-LANG, by integrating Dynamic and Diagnostic assessment, aims to **provide a fuller and more individualised diagnosis** of learner L2 development
 - Learner responsiveness to **mediation** is essential for understanding how much additional support they likely require to develop
- By **combining AI-enhanced Revita exercises with teacher assessments** in classroom, we get new diagnostic insights and opportunities for supporting learner development
- **Next steps:** piloting online Revita exercises with mediation & starting to them larger scale in several schools in the autumn
 - Including new workshops with the teachers



Thank you for your attention!

For more information:

https://r.jyu.fi/DDLANG_en

ari.huhta@jyu.fi



References

- Alderson, J. C., Brunfaut, T., & Harding, L. 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.
- Alderson, J C, Haapakangas, E-L, Huhta, A, Nieminen, L & Ullakonoja, R. 2015. *The Diagnosis of Reading in a Second or Foreign Language*. London: Routledge
- Black, P. & Wiliam, D. 1998. Assessment and classroom learning. *Assessment in Education*, 5(1), 7–71.
- Bloom, B., Hastings, J. T., & Madaus, G. 1971. *Handbook on Formative and Summative Evaluation of Student Learning*. McGraw-Hill.
- Davies, A. 2014. Fitty years of language assessment. In A.Kunnan (ed). The Companion to Language Assessment. Wiley.
- Hamp-Lyons, L. & Lynch, B. 1998. Perspectives on validity. A historical analysis of language testing conference abstracts. In A.Kunnan (ed.) Validation in Language Assessment. Routledge.
- Harding, L., Alderson, J. C., & Brunfaut, T. 2015. Diagnostic assessment of reading and listening in a second or foreign language: Elaborating on diagnostic principles. *Language Testing*, 32(3), 317–336.
- Huhta, A. 2008. Diagnostic and formative assessment. In Spolsky, B & Hult, F (eds.) *Handbook of Educational Linguistics*. Blackwell. 469-482.
- Huhta, A. 2023. Improving the impact of technology on diagnostic language assessment. In Sadeghi, Karim & Douglas, Dan (Eds.). Fundamental Considerations in Technology Mediated Language Assessment. Cambridge University Press.
- Huhta, A., Harsch, C., Leontjev, D & Nieminen, L 2024. *The Diagnosis of Writing in a Second or Foreign Language*. Routledge.
- Jang, E. 2009. Cognitive diagnostic assessment of L2 reading comprehension ability: Validity arguments for Fusion Model application to *LanguEdge* assessment. *Language Testing*, 26(1), 31–73.
- Jang, E. & Wagner, M. 2014. Diagnostic feedback in language classroom. In A. Kunnan (Ed.), *Companion to language assessment*. Wiley-Blackwell.
- Laveault, D., & Allal, L. 2016. Assessment for learning: Meeting the challenge of implementation. Springer.



- Leontjev, D. 2016. *ICAnDoiT: The impact of computerised adaptive corrective feedback on L2 English learners* [Doctoral dissertation, University of Jyväskylä] <https://jyx.jyu.fi/handle/123456789/49346>
- Leung, C. & Mohan, B. 2004. Teacher formative assessment and talk in classroom contexts: Assessment as discourse and assessment of discourse. *Language Testing*, 21(3), 335–359.
- Mei, H. & Chen, H. 2022. Cognitive diagnosis in language assessment: A thematic review. *RELC Journal*. Online first.
- Lantolf, J., & Poehner, M. 2004. Dynamic assessment of L2 development: Bringing the past into the future. *Journal of Applied Linguistics*, 1(2), 49–72.
- Poehner, M., & Lantolf, J. 2013. Bringing the ZPD into the equation: Capturing L2 development during Computerized Dynamic Assessment (C-DA). *Language Teaching Research*, 17(3), 323–342.
- Poehner, M. & Leontjev, D. 2020. To correct or to cooperate: Mediation processes in L2 development. *Language Teaching Research* 24 (3), 295-316.
- Purpura, J. & Turner, C. 2023. *Learning-Oriented Assessment in Language Classrooms: Using Assessment to Gauge and Promote Language Learning*. Routledge
- Rea-Dickins, P. & Gardner, S. 2000.. Snares and silver bullets: Disentangling the construct of formative assessment. *Language Testing*, 17(2), 215–243.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119–144..
- Scriven, M. 1967. The methodology of evaluation. In R.Tyler, R.Gagné, & M.Scriven (eds.), *Perspectives of Curriculum Evaluation* (pp. 39–83). Rand McNally.
- Turner, C. & Purpura, J. 2016. Learning-oriented assessment in second and foreign language classrooms. In Tsagari, D & Banerjee, J. (eds) *Handbook of Second Language Assessment*. De Gruyter Mouton, pp. 255-274
- William, D. 2011. *Embedded formative assessment*. Solution Tree Press.
- Vygotsky, L. 1978. *Mind in society. The development of higher psychological processes*. Harvard University