

# DEVELOPING A TASK SPECIFICATION FOR WRITTEN INTRALINGUISTIC MEDIATION

**Caroline Shackleton**

**Nathan Turner**

**Adolfo Sánchez Cuadrado**

**7th May 2026**



**API EDUCATION  
FORUM**



**UNIVERSIDAD  
DE GRANADA**



**CENTRO DE  
LENGUAS  
MODERNAS**

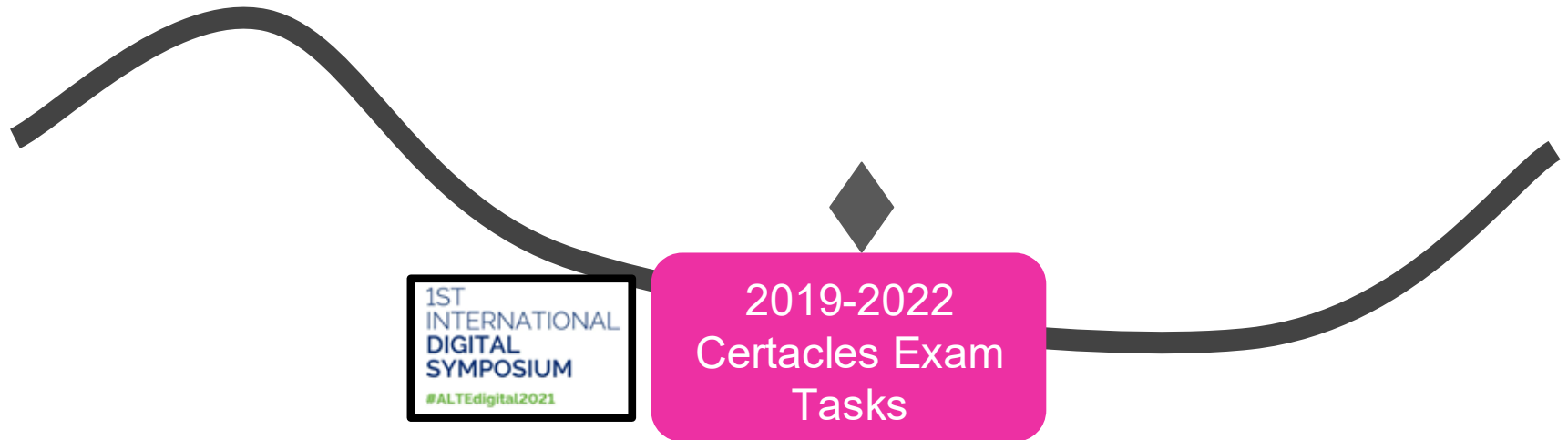
## Inclusion of mediation in assessment practices



Assessment of intralinguistic mediation. Integration in a Certaces exam.

Caroline Shackleton, Nathan Turner and Adolfo Sánchez Cuadrado

## Inclusion of mediation in assessment practices



1ST  
INTERNATIONAL  
DIGITAL  
SYMPOSIUM  
#ALTEdigital2021

2019-2022  
Certacles Exam  
Tasks

# Context

- *CertAcles* accreditation exams
- CEFR-related
- Majority of test takers between 18 and 25 (mobility programmes and degree accreditation)
- Lingua Franca (intra-linguistic mediation)
- Attempt to expand the test construct – plan for positive washback

# Evidence-centred design

Layers:

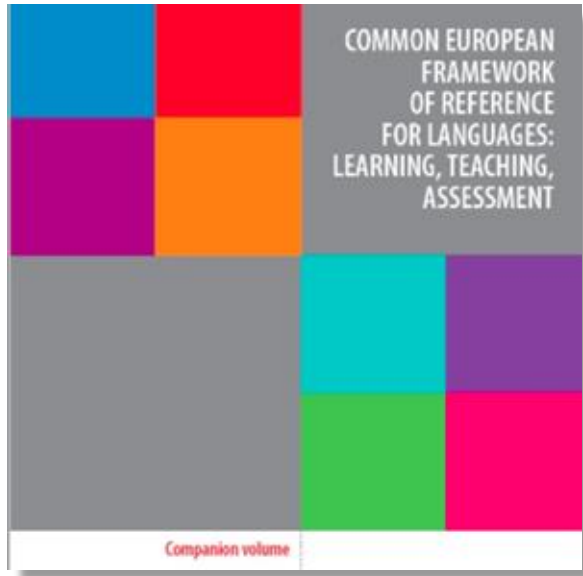
1. Domain analysis (job analysis)
2. Domain modelling

Claims we want to make about test-takers and the necessary observable evidence to support those claims.

## The 'thinking stage' of ECD (McNamara & Roever, 2006)

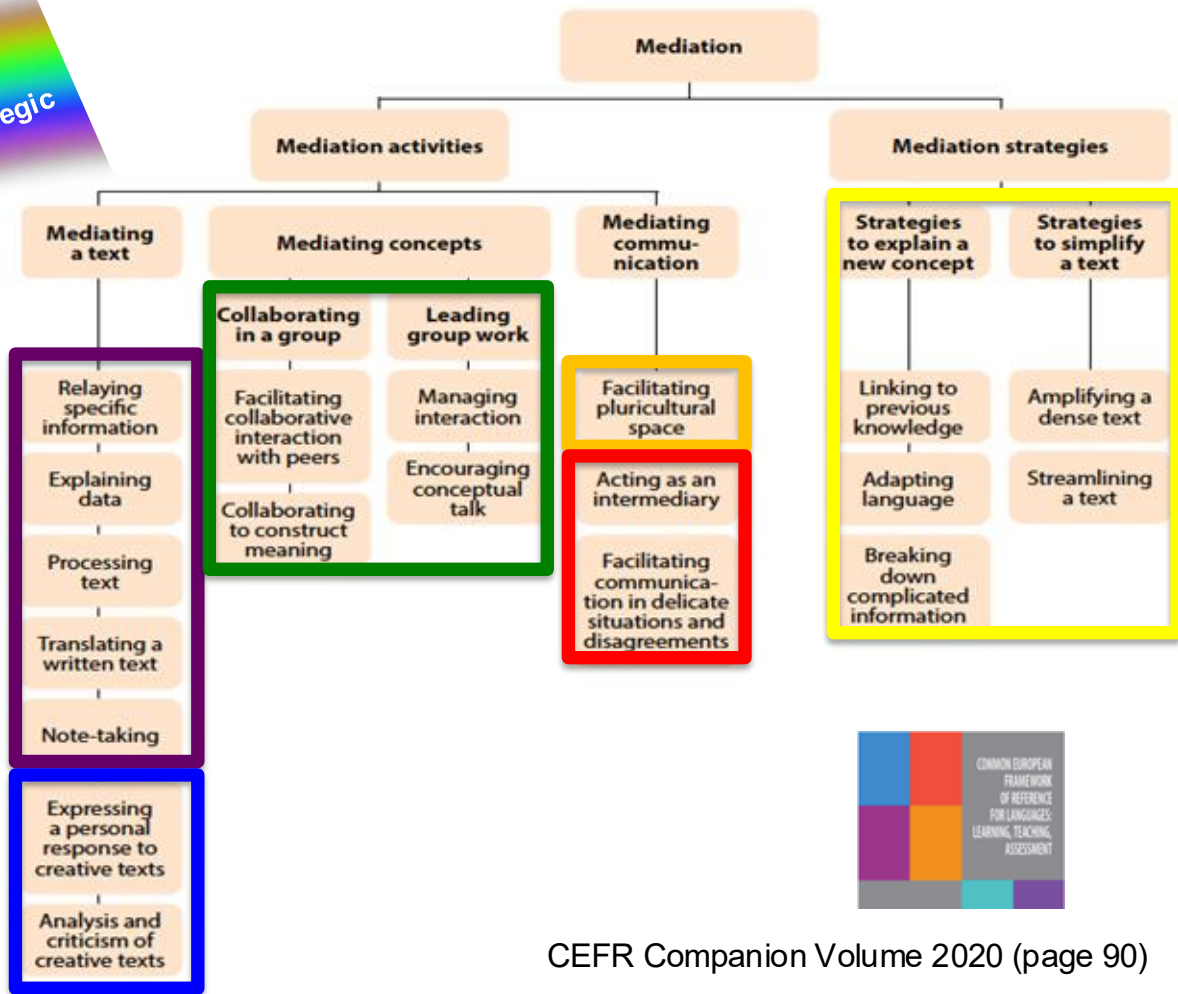
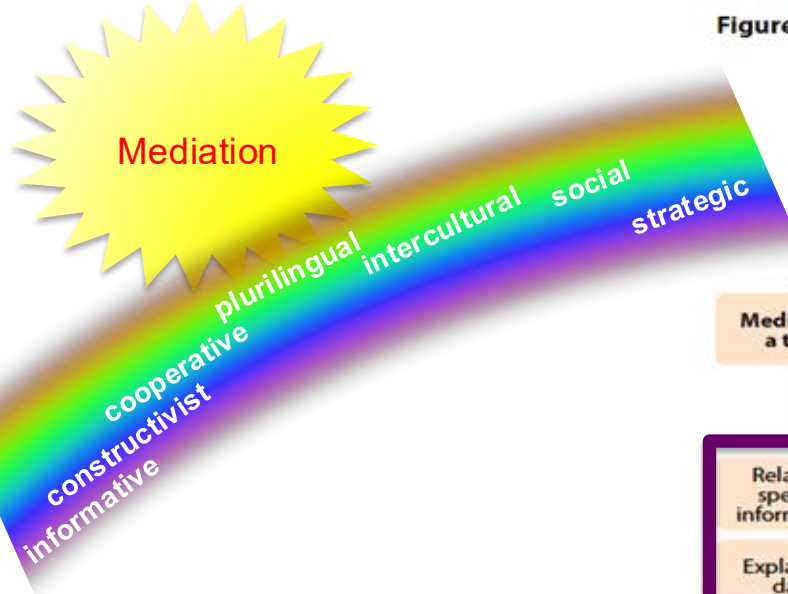


# Domain analysis



- What knowledge, skills and abilities (KSAs) are important for the construct of mediation?
- Rich description of the construct:

Figure 14 – Mediation activities and strategies



# List of KSAs

## **FOR LINKING TO PREVIOUS KNOWLEDGE:**

Encouraging people to activate prior knowledge.  
Providing extended definitions and explanations.  
(Help) To make connections to previous knowledge and experiences.  
Comparing a new concept to one people are already familiar with.  
Giving examples related to people's everyday experiences.

## **FOR AMPLIFYING A DENSE TEXT:**

Elaborating and exemplifying.  
Adding helpful detail.  
Recapitulating and repeating the main points.  
Adding illustrations.  
Using repetition and redundancy.  
Modifying style.

## **FOR ADAPTING THE LANGUAGE:**

Paraphrasing  
Adapting speech/delivery  
Explaining technical terminology

## **FOR STREAMLINING A TEXT:**

Removing unnecessary sections.  
Excluding non-relevant or repetitive information.  
Taking into account the intended audience.  
Highlighting key information.

## **FOR BREAKING DOWN COMPLICATED INFORMATION:**

Presenting logical order of information.  
Repeating key aspects.  
Listing the most important aspects/steps.

Figure 14 – Mediation activities and strategies

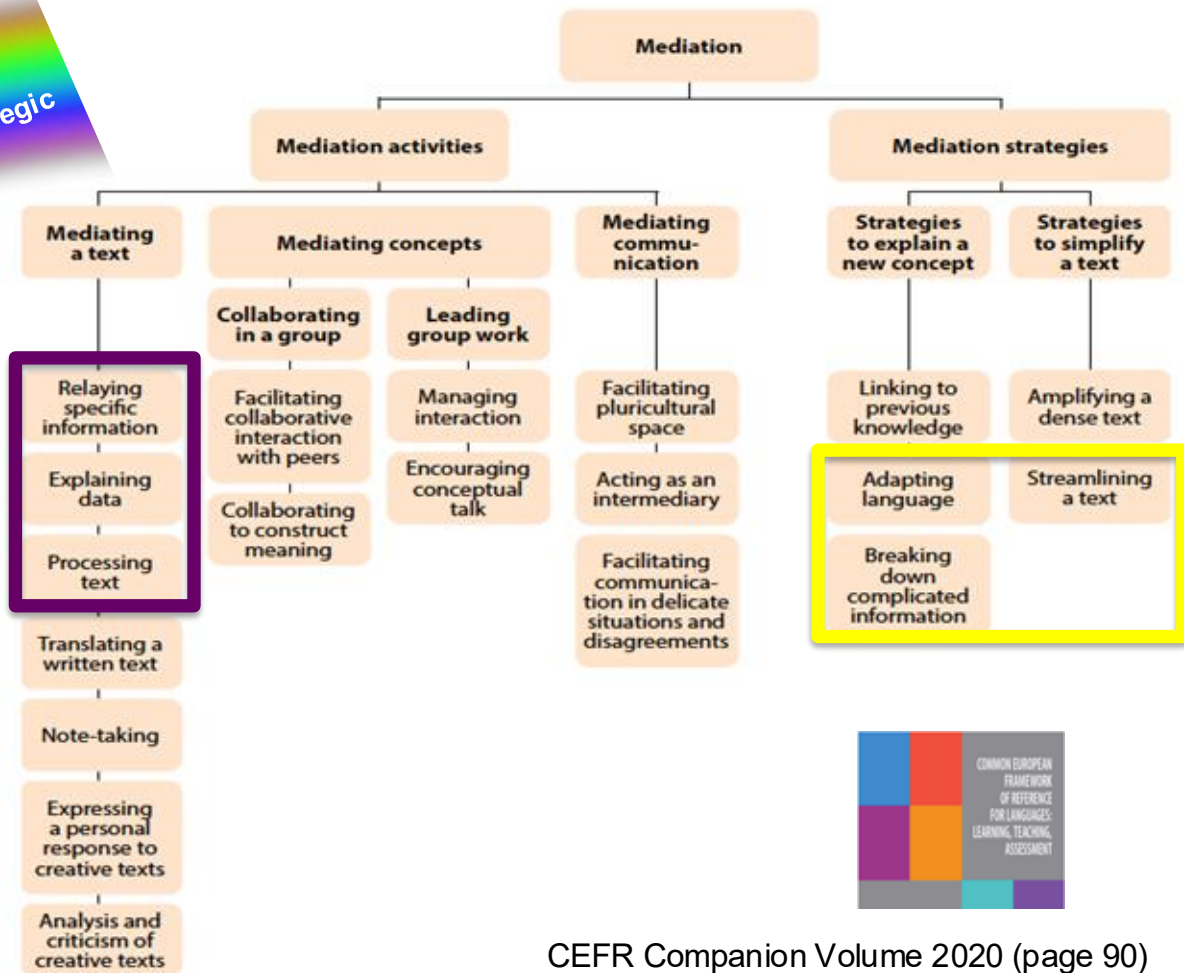
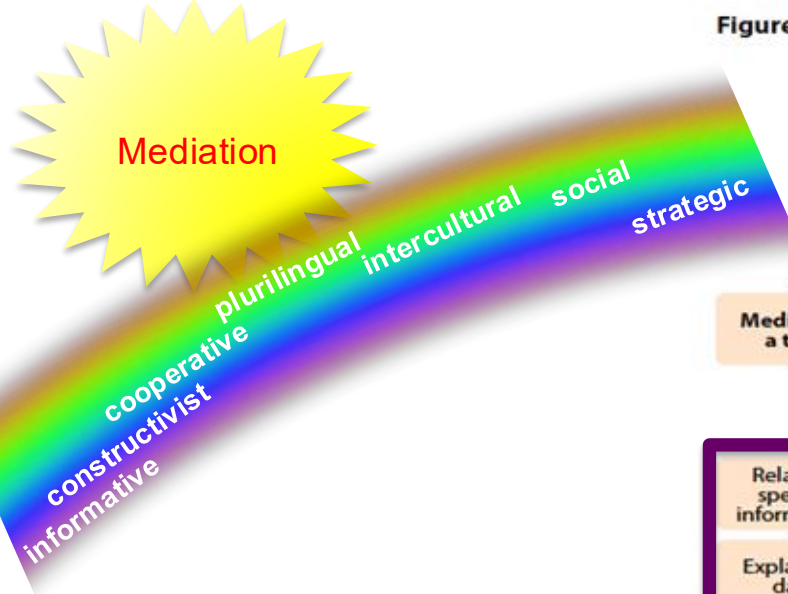
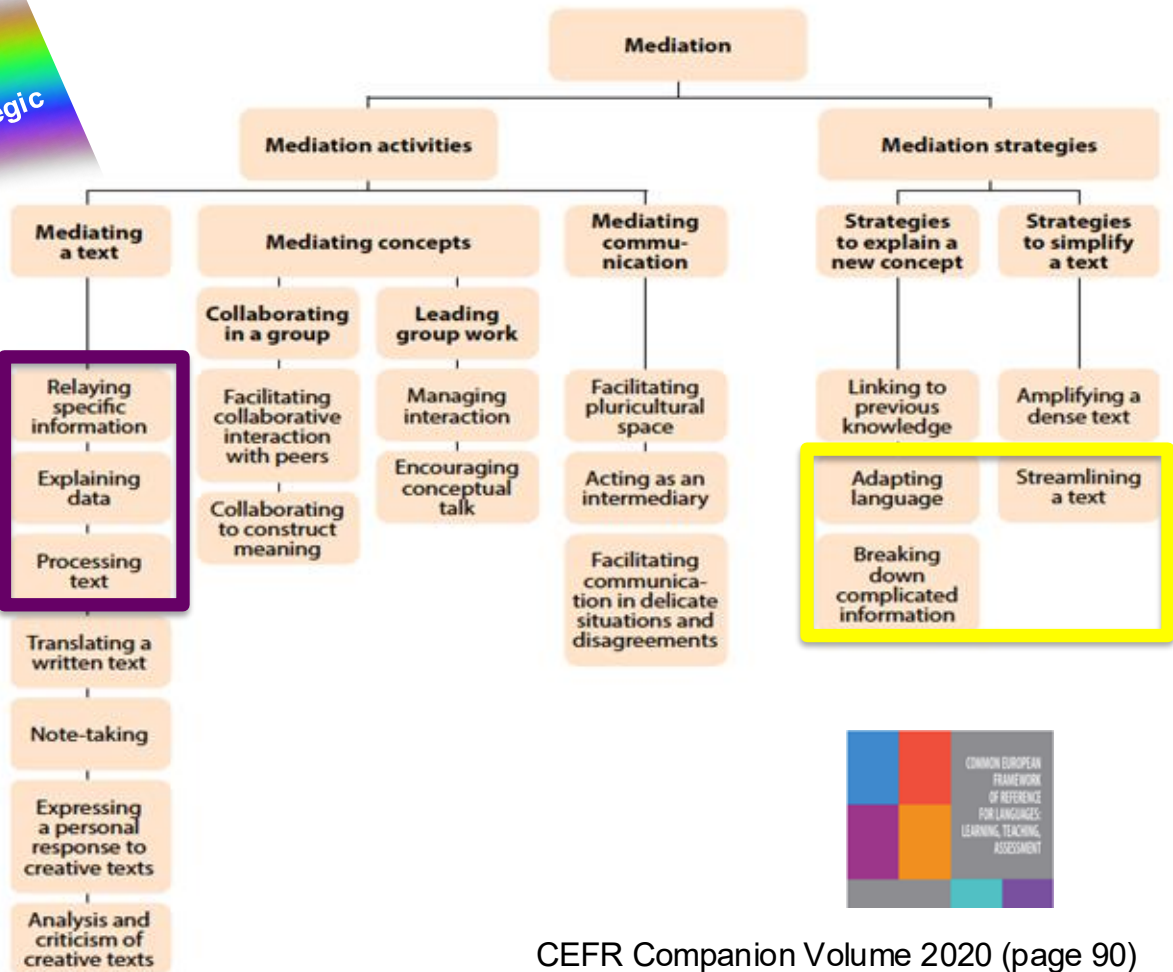


Figure 14 – Mediation activities and strategies

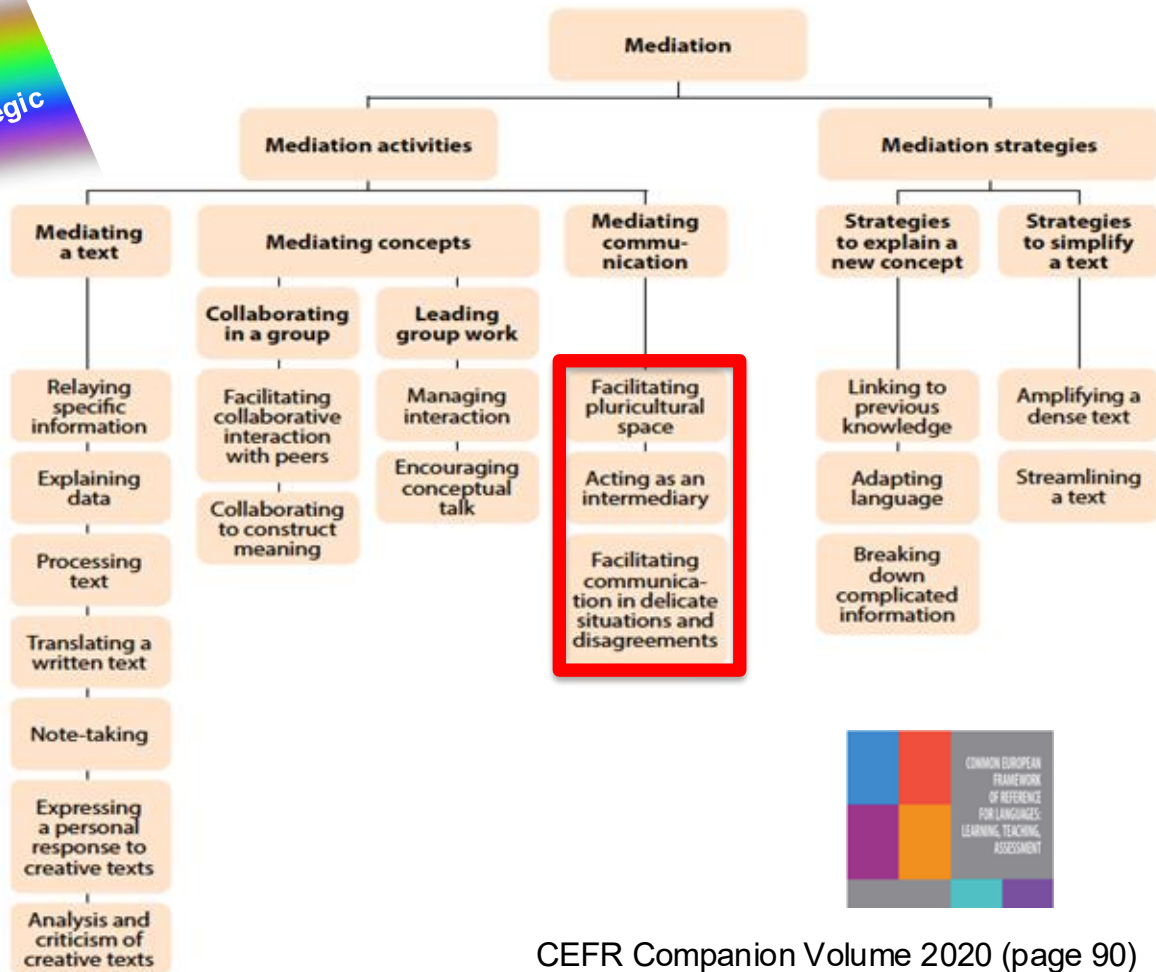
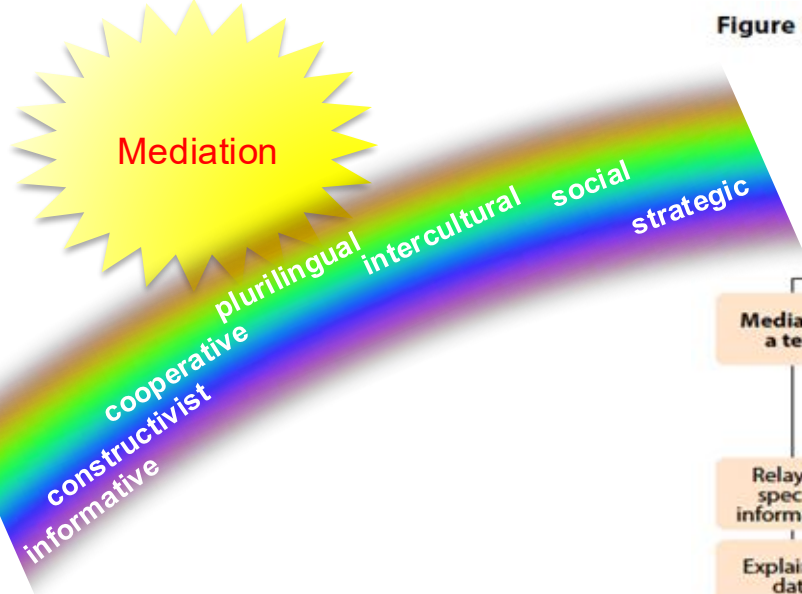


### Tasks for mediating texts

- + Focus on reprocessing of meaning.
- + No too demanding to design.
- + Strategies provided by CEFR CV.
- + Easy to relate to previous testing practices.
- Limited focus.



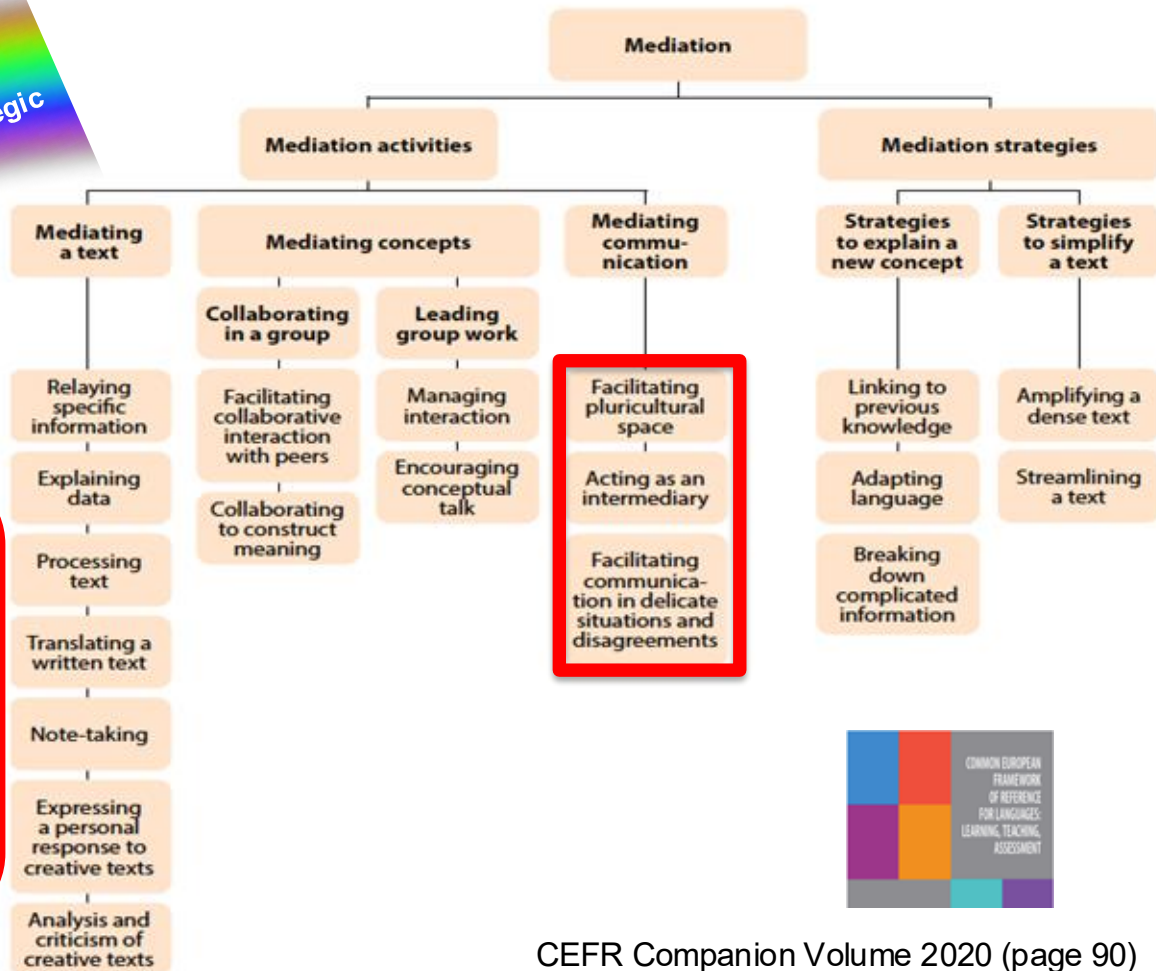
Figure 14 – Mediation activities and strategies



# Mediating communication

1. Facilitating pluricultural space
1. Acting as intermediary in informal situations (with friends and colleagues)
1. Facilitating communication in delicate situations & disagreements (**written response**)

Figure 14 – Mediation activities and strategies



**Tasks for mediating communication**

- + Focus on the social dimension of candidates.
- + Construct expanded: facilitating relations when living together.
- + Easy to relate to everyday communicative situations.
- Need to detect strategies (KSAs).



# List of important KSAs

## FOR MEDIATING COMMUNICATION

- (Help) To identify sociocultural and sociolinguistic differences
- (Help) To identify and manage ambiguity.
- Discussing similarities and differences.
- Getting clarification on what is meant.
- Explaining cultural references.
- Using follow-up questions for the sake of clarification.
- Giving additional explanations if necessary.
- Preparing beforehand for the communicative task/situation.
- Pausing to plan utterances.
- Asking for repetition and clarification.
- Attenuating and using diplomatic language.
- (Help) To identify common ground.
- Posing neutral questions.
- Highlighting possible areas of agreement.

# Domain modelling

Facilitating communication in delicate situations & disagreements (p. 117)

## B2 Claim:

Can, by asking questions, **identify areas of common ground** and invite each side to **highlight possible solutions**.

Can **outline the main points** in a **disagreement** with **reasonable precision** and **explain the positions of the parties** involved.

Can **summarise** the statements made by the two sides, **highlighting areas of agreement** and **obstacles to agreement**.

# Domain modelling

Facilitating communication in delicate situations & disagreements (p. 117)

## B1 Claim:

Can ask parties in a disagreement **to explain their point view**, and can **respond briefly to their explanations**, provided the topic is familiar and the parties express themselves clearly.

Can demonstrate their **understanding of the key issues in a disagreement on a topic familiar** to them and make **simple requests for confirmation and/or clarification**.

# Sub claims (KSAs) – prioritised

- Ability to take into account the intended audience
- Ability to (help to) identify common ground
- Ability to highlight possible areas of agreement
- Ability to highlight possible obstacles to agreement
- Ability to attenuate and use diplomatic language
- Ability to give additional explanations if necessary
- Ability to ask for repetition and clarification
- Ability to summarise/paraphrase
- Ability to discuss similarities and differences
- Ability to pose neutral questions
- Ability to highlight key information
- Ability to get clarification on what is meant

# Domains in context – Personal

**Context:** University/Erasmus flat share.

**Situation:** Disagreement/misunderstanding over payment (how and when rent due, how and when bills are paid, services available, etc.), damage in a flat (whose fault, who should pay, how much deposit should be returned, etc.), rules to follow, etc.

## Example task

- The candidate receives a series of SMS messages from flatmate about a broken water heater that the landlord refuses to pay for as well as an email from the landlord threatening eviction.
- The candidate is instructed to write to both parties in order to help resolve the conflict.

# Evidence 1

“Cycles of iteration and refinement both within and across layers are expected and appropriate.” Mislevy and Haertel (2006, p.7)

## Focus group 1

- “We have the same opinion and the landlord is the opposite so it’s difficult.”
- “I prefer to write to my friend and the landlord, not both.”
- “It would be more logical to speak to my flatmate or send a message saying we’ll speak later or write to the landlord and copy him in.”
- This was the most difficult part .. I don’t know how I can write to both.”

## Focus group 2

- “To me the task seems very real because this could happen to everybody.”

# Evidence 2

'Think alouds' – Macro planning (n=7)

## Salient themes:

1. Role of the mediator
2. Positioning
3. Drawing on prior knowledge
4. Additional explanations/justifications

'Assessment scales need to give important KSAs priority'

Initial piloting – observable evidence (n=24)

KSA	Total/24
Take into account the intended audience	23
<b>FACILITATE (CONTINUED) COMMUNICATION</b>	21
Attenuate and use diplomatic language	21
Give additional explanations if necessary	21
Highlight possible areas of agreement/ <b>SOLUTIONS (SUGGEST A COMPROMISE)</b>	17
Highlight key information	17
Summarise/paraphrase	10
Pose neutral questions/ <b>MAKE REQUESTS</b>	10
Highlight possible obstacles to agreement	8
Get clarification on what is meant	7
Identify common ground	3
Explain cultural references	2
Discuss similarities and differences	0

# Task shell – Written response

## **Mediating communication 1: Resolving/avoiding conflict/misunderstanding**

**Inputs A & B:** Two written texts of approximately 100 words each (e.g.,. formal/informal email, SMS message, social media post/message). Care must be taken to avoid bias in favour of one party.

**Input C:** Context statement and task instructions (max. 100 words).

**Response mode:** Written. Candidate must write to one of the parties independently or to both parties depending on the context, in the required format and style in order to mediate between them.

**Topic areas:** Personal, public and educational domains relevant to University undergraduates from any faculty (see CV p.240 - 241).

## Task 1

### Input text A

On a trip back home to see your family, you get the following SMS messages from your Erasmus flatmate:

Listen, we've got problems. The water heater has stopped working and there's no hot water! I rang the owner but he insists that we must have broken it and refuses to do anything...

I was under the impression that fixed items in the flat are the landlord's responsibility!

If I can't get a decent response from him by the end of today, I'm just going to call a service technician and take the cost out of the next rent!

I'm afraid I got angry and shouted at him

I'm just really sick and tired of his attitude

Check your email — now he's threatening to kick us out because I haven't paid my part of the rent!

But he hasn't even come and looked at the problem with the TV we called him about two weeks ago.

I'm not paying till the problems are sorted!

Input text B

**When you check your inbox, you find an email from your landlord:**

Dear tenants,

I am writing to you to inform you of my intention to cancel your accommodation contract next month after today's horrible argument. I am tired of your bad attitude and the way you mistreat the flat and break its contents. I remind you that it is not my responsibility to pay for any items that you have broken. Also, you have now been late with the rent and bills on a number of occasions. Enough is enough. I also expect you to pay for the broken TV and water heater.

Regards,

L. Hancock

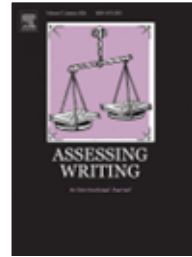
## Input text C

You have been living in your current flat for two years and used to have a good relationship with the owner of the flat but since your friend moved in six months ago you have been having problems—your friend is argumentative and often pays late. The situation is quite serious, you like your current flat and you don't want to have to look for another one. You still have a few days left at home with your family, so you'll have to get in touch by email. Write an email to the owner to see if you can help to resolve the situation

### **In the email you should:**

- Apologize for and justify your friend's behaviour
- Ask the owner to explain his reasons and position
- Suggest ways of reaching a compromise and improving their relationship

## Inclusion of mediation in assessment practices



2023-2025  
Certacles Exam  
Rubrics

# DEVELOPING A RATING SCALE FOR WRITTEN INTRALINGUISTIC MEDIATION

**Caroline Shackleton**

**Nathan Turner**

**Adolfo Sánchez Cuadrado**

**7th May 2026**



**API EDUCATION  
FORUM**



**UNIVERSIDAD  
DE GRANADA**



**CENTRO DE  
LENGUAS  
MODERNAS**

# Our Context

- University of Granada's Modern Language Center
- Large European university with a high proportion of international students
- Bi-level CertAcles test of English at CEFR B1 and B2 levels administered three times a year as part of UGR internationalization policies—free of charge for UGR students



Access Erasmus-type mobility programs



Degree completion requirement



Further academic study

# English as a Lingua Franca Mediation Tool

---

## Plurilingual & Pluricultural Contexts

English is used as a lingua franca mediation tool in these complex environments.

## Navigating Unfamiliar Contexts

- Users must adapt, accommodate, and use interpersonal strategies (Harding, 2014; Harding & McNamara, 2018).
- Proficiency in pragmatics outweighs proficiency in grammar (Canagarajah, 2006).

## Intralinguistic Mediation

Reflects these behaviours and is extremely relevant to our context.

---

# Test re-development project: Overview



## Reception

- Reading
- Listening



## WPIM

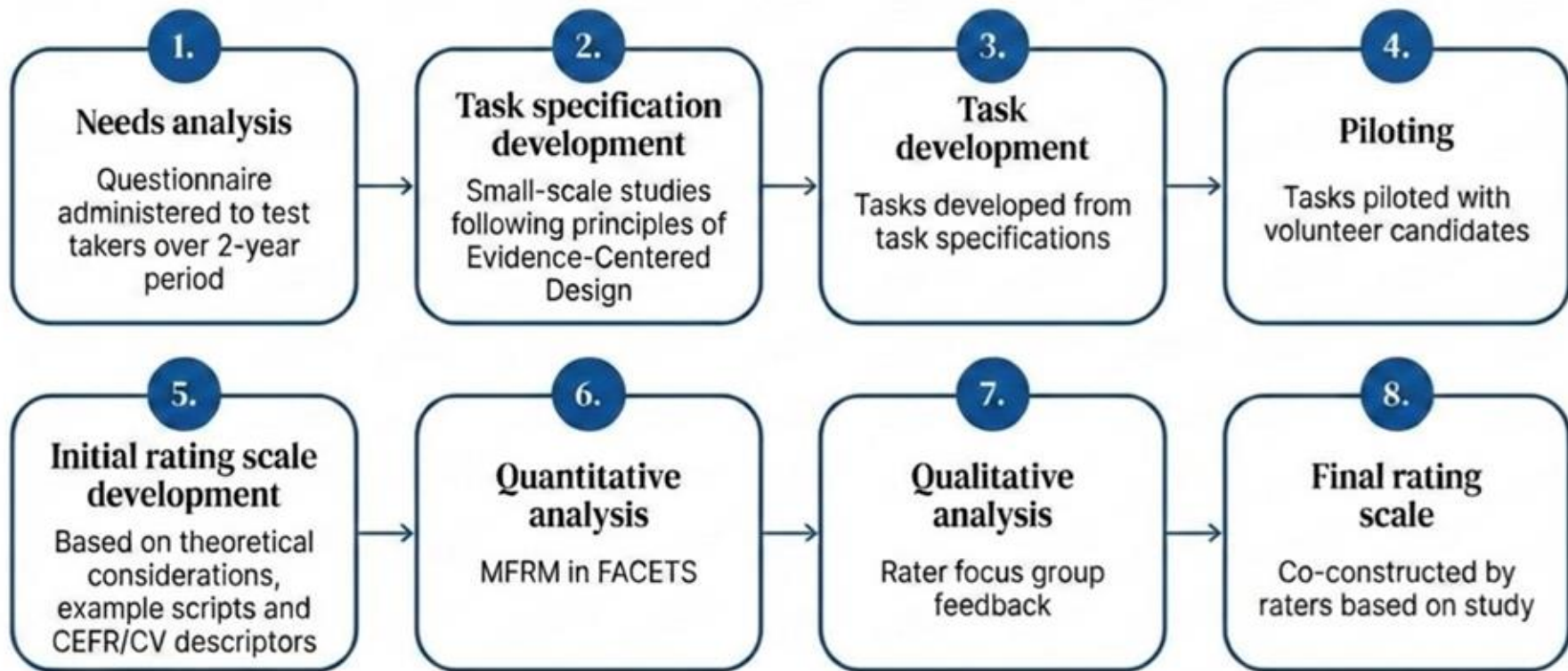
1. Listening into writing
2. **Reading into writing**



## SPIM

1. Presentation
2. Group task

# Project Design



# Written Production, Interaction and Mediation Paper

1.

## **Needs analysis**

Questionnaire  
administered to test  
takers over 2-year  
period

**Emails form of written  
communication most used**

## Task 2: Reading into writing task with an email as output

2.

### Task specification development

→ Small-scale studies following principles of Evidence-Centered Design

**1. Disagreements**

2. Acting as an intermediary

3. Mediating text and data

# Task specification 1: Mediating communication: facilitating communication in delicate situations and disagreements

3.

## Task development

Tasks developed from task specifications

1. Personal: a disagreement between landlord and tenants in the context of an Erasmus flat share.
2. Educational: a disagreement between classmates in the context of class group work.
3. Public/Academic: a disagreement between a student and a university teacher in the context of a student union body.

4.

## Piloting

Tasks piloted with  
volunteer candidates

**Volunteers from language center (52 students)**

**Mainly B1 and B2**

**80 scripts collected**

# Background

## CEFR/CV context

The CEFR/CV framework is described mainly in terms of implementation in the classroom, a situation viewed as a real social context (North, 2021, 2022).

## Formative assessment focus

Previous attempts to incorporate mediation into assessment have mostly taken place in the realm of formative assessment, where process-oriented evaluation is prioritized over scores (e.g., Gadomska, 2024; Saito, 2020; Sanchez Cuadrado, Shackleton & Turner, 2022 ; Stathopoulou, 2025).

The nature of the calibrated CEFR/CV descriptors for mediation seems to lend itself better to task design, portfolios, and classroom assessment as opposed to proficiency exams.

# Rating Scale Development

## 1. Mixed Approach



Based on integrating test-external sources and test-internal sources (actual performances and tasks)



## 2. Score Allocation Process



Holistic / Analytic



Checklists: Focusing on salient features

## Insights from other relevant approaches to writing assessment

### 1. Integrated reading into writing.

Analytic scales

Source use criteria

Context and audience engagement paramount for a mediation activity

### 2. Task-based language assessment and task fulfillment

Task fulfillment is a holistic dimension

Functional adequacy – shift towards proficiency in pragmatics

# Initial rating scale development



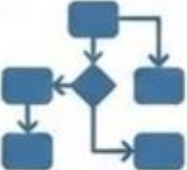
**Content**



**Context**



**Comprehensibility**



**Organization**

## Descriptors for each band



**Raters worked  
together**



**Scripts ranked -  
salient features  
identified**



**CEFR/CV  
descriptors  
consulted**

Level	CONTEXTUAL APPROPRIATENESS  (for communicative purpose – register, layout, style, socio-pragmatics)	CONTENT  (relevance, accuracy and transformation of ideas to create new meaning)	ORGANIZATION  (connecting and organizing relevant mediated ideas)	COMPREHENSIBILITY  (amount of effort and re-reading required)
7. C1	<p>Consistently controls register, style, and layout.</p> <p>Uses hedging, softening, and persuasive language.</p> <p>Maintains a diplomatic tone throughout, evidencing sensitivity to all parties' opinions.</p>	<p>All key points are fully addressed and precisely explained and interpreted.</p> <p>Conveys clearly and accurately the significant ideas using own language.</p> <p>Includes nuances and subtleties.</p> <p>Suggests relevant constructive next steps or compromises and solutions.</p>	<p>Demonstrates sophisticated organization of ideas, including complex structures, thematic progression, and effective use of discourse markers.</p> <p>Ideas flow logically and persuasively.</p>	<p>Text is accessible and engaging on first reading.</p> <p>Highly comprehensible, with good grammatical accuracy and wide lexical range.</p> <p>Errors are rare and do not affect clarity.</p>
6. B2+				
5. B2	<p>Adapts register, style, and layout effectively to suit communicative purpose and audience most of the time.</p> <p>Attempts attenuation and polite language to soften conflict and encourage dialogue.</p>	<p>All key points are addressed and significant ideas/arguments have been identified and are reliably summarized.</p> <p>Information has been well-paraphrased with little evidence of lifting.</p> <p>There is evidence of additional</p>	<p>Organizes ideas coherently and cohesively.</p> <p>Uses a variety of connectors and cohesive devices to show relationships between ideas effectively.</p>	<p>Clear and fluent communication with minimal effort required on behalf of the reader. Text is easy to follow on first reading.</p> <p>Uses some complex forms. Errors do not hinder understanding.</p> <p>Vocabulary range is sufficient</p>

## The study

**RQ1.** To what extent do the raters apply the analytic rating scale consistently; does the scale reliably separate the candidates?

**RQ2.** To what extent do the four dimensions of the rating scale reliably capture overall task fulfillment for this task specification?

**RQ3.** Do raters think that any changes should be made to the rating scale based on the quantitative results and their experience during a pilot rating session?

# Quantitative Analysis

Judging plan is a fully-crossed design (all raters marked all scripts)

MFRM in FACETS Results

Rasch hybrid partial credit model was used, which pays particular attention to the functioning of an analytic scale and its categories.



## Rater Consistency

Raters used scale in a consistent manner.



## Scale Reliability

Analytic scale shown to work well and reliably separate candidates into different proficiency levels.



## Scale Structure

Scale is unidimensional and each of the four dimensions contributes to the overall score allocated to the task response.



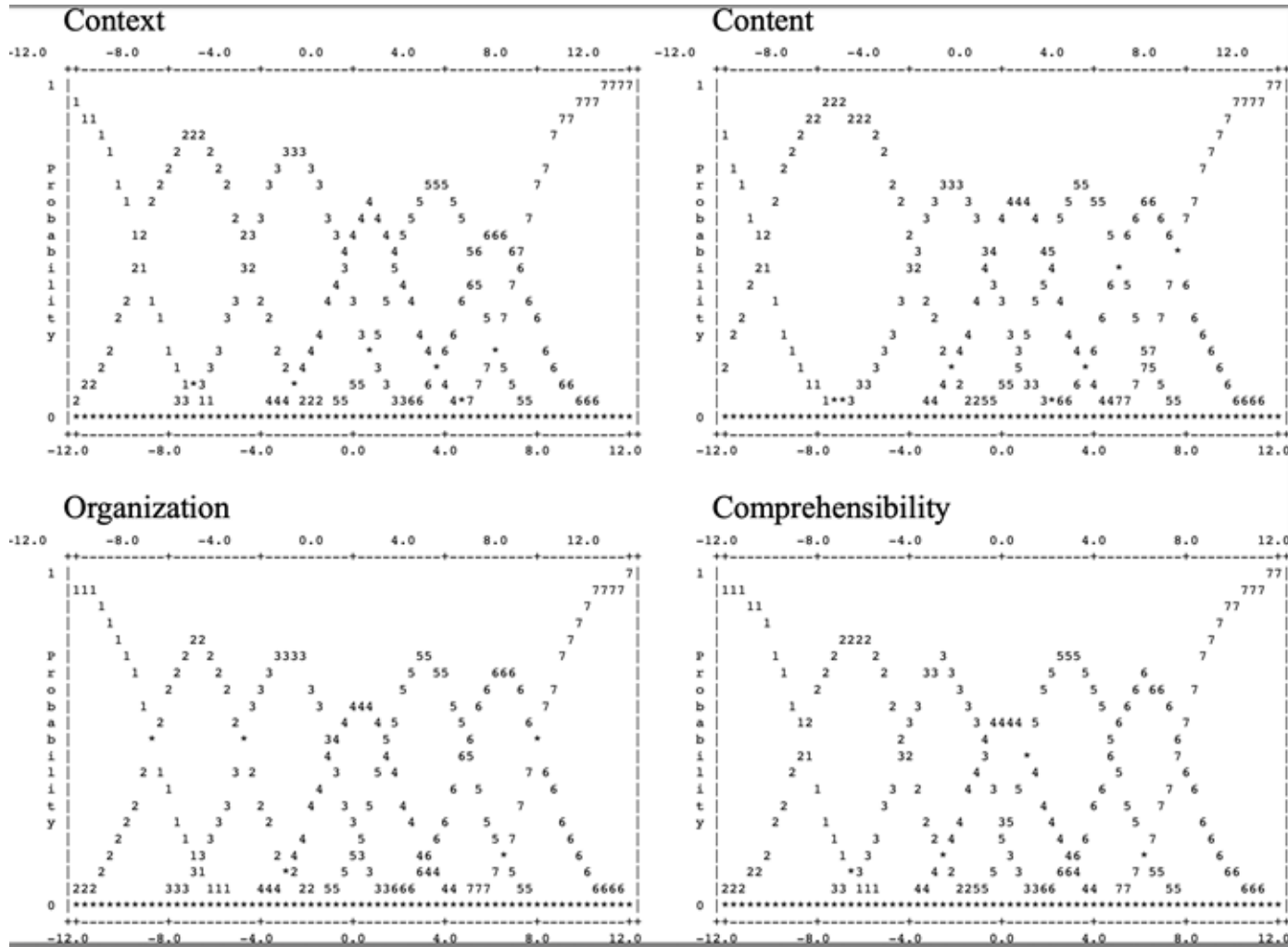
## Step Structure

Step structure for each of the four dimensions is similar so data re-analyzed using rating scale model.

## Linacre's (2002) guidelines for rating scale functioning:

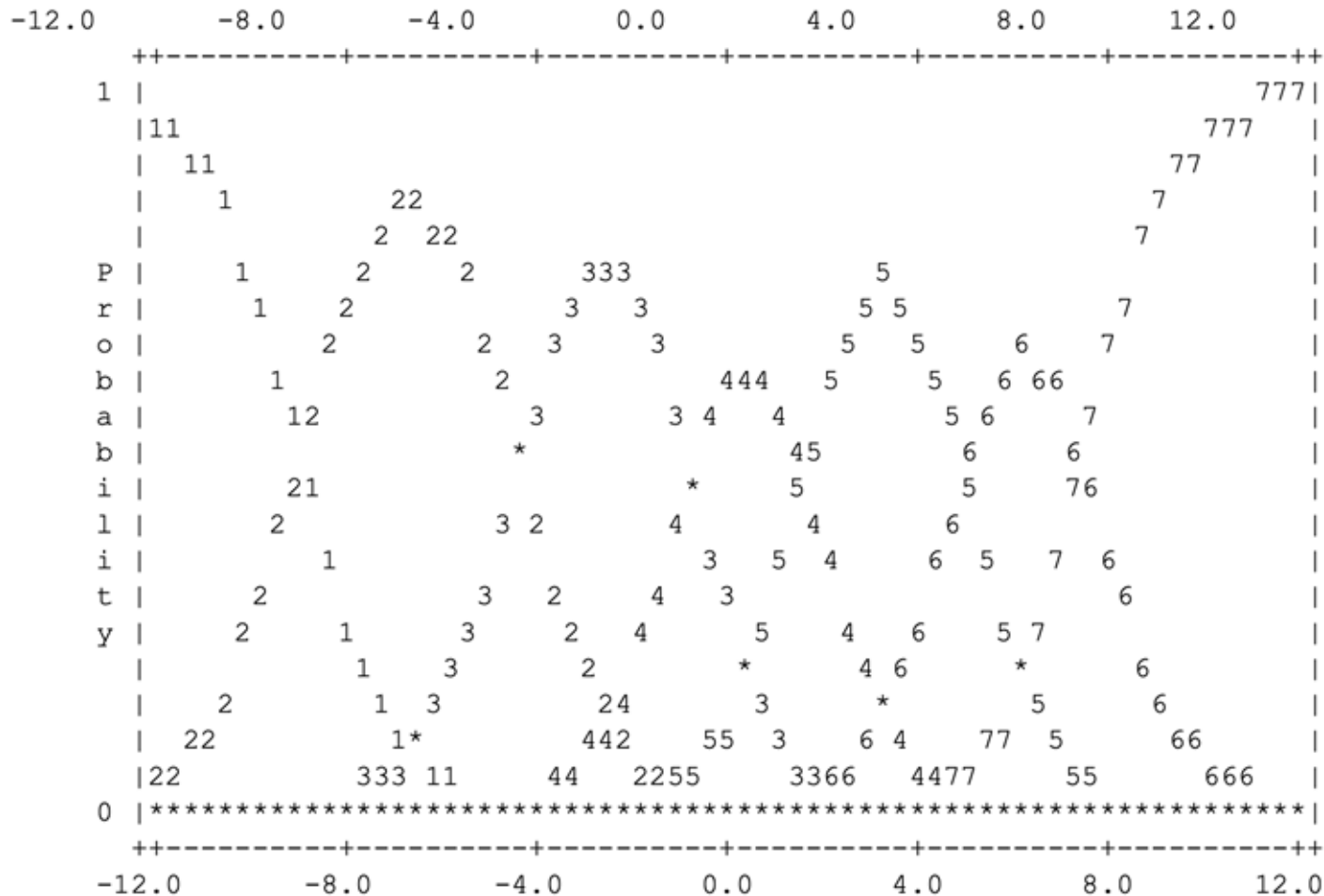
1. There should be more than 10 observations per category.
2. The frequency of data points in each category should result in a smooth distribution, i.e. the distribution should not be jagged.
3. Average measures should not be disordered. They should advance monotonically.
4. Average measures should be near their expected values.
5. Outfit Mean-square values should be less than 2.
6. Rasch-Andrich thresholds (step difficulty, step calibration) should increase at each score level by 1.4 or more, but no more than 5 logits.

# Probability curves for category rating scale structures (PCM)



# Vertical ruler PCM analysis

Measr	-rater	+candidate	-task	-Criteria	S.1	S.2	S.3	S.4					
12	+	4	+	+	+	(7)	+	(7)	+	(7)	+	(7)	
11	+	10	+	+	+	+	+	+	+	+	+	+	
10	+		+	+	+	+	+	+	+	+	+	+	
9	+		+	+	+	+	+	+	+	+	+	+	
8	+	1	+	+	+	+	---	+	---	+	---	---	
7	+		+	+	+	+	+	+	+	+	+	+	
6	+	32	+	+	+	6	+	6	+	6	+	6	
5	+	31 43	+	+	+	---	+	---	+	---	+	---	
		34 36 40											
4	+	8	+	+	+	+	+	+	+	+	+	+	
		35				5	+	5	+	+	+	+	
3	+		+	+	+	+	+	+	5	+	5	+	
		45 46											
2	+		+	+	+	---	+	---	+	+	+	+	
		37 38 6							---	---	---	---	
1	+	3	+	+	+	+	+	+	+	+	+	+	
		27 39 47 48				4	+	4	+	+	+	+	
		50											
*	0	* 1	* 17 25 33 44 5	* 2 3 *	*	*	*	4	*	4	*	4	
		2 4	11 19 49	1		---	---	---	---	---	---	---	
-1	+		14 28	+	+	+	+	+	+	+	+	+	
			12 2 41										
-2	+		24 42	+	+	+	3	+	+	+	+	+	
			16 18 22 23 30			3	+	+	3	+	3	+	
-3	+		13 7	+	+	+	+	+	3	+	+	+	
			3										
-4	+		20 21 51	+	+	+	---	+	+	+	+	+	
			15			---	---	---	---	---	---	---	
-5	+		29	+	+	+	+	+	---	+	+	+	
			52										
-6	+			+	+	+	+	+	+	+	+	+	
-7	+			+	+	+	2	+	2	+	+	+	
-8	+			+	+	+	+	+	+	+	+	+	
										---	---	---	
-9	+			+	+	+	+	+	+	+	+	+	
			26			---	---	---	---	---	---	---	
-10	+			+	+	+	(1)	+	(1)	+	(1)	+	(1)
Measr	-rater	+candidate	-task	-Criteria	S.1	S.2	S.3	S.4					



# Notable findings (potential problems)

- Only 2 marks with standardized residuals in the  $\pm 3$  range
- Five misfitting candidates with Infit Mean-square higher than 1.50 (9.6%)
- Little evidence of systematic bias of raters across criteria - the result is not significant ( $\chi^2 = 16.6$ ,  $df = 16$ ,  $p = 0.41$ ). Only one statistically significant interaction

# Qualitative Analysis



## 1. Rater Feedback

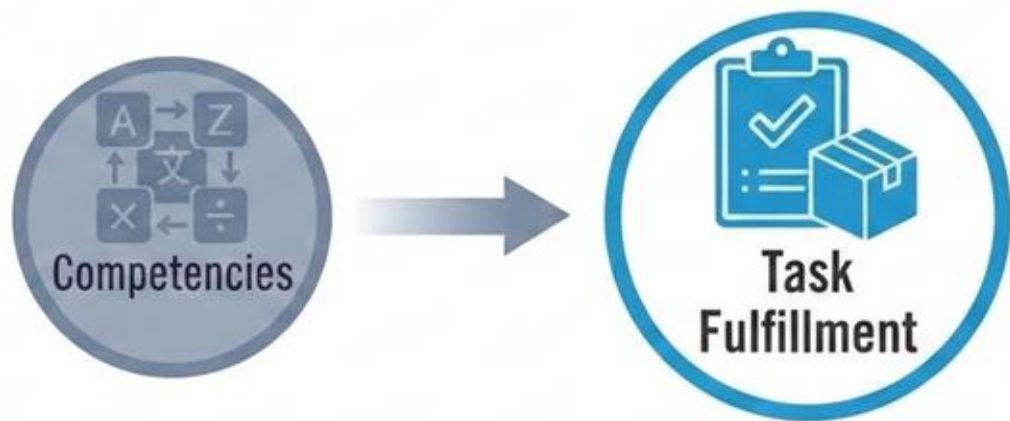
- Raters think the scale is practical and useable
- A 'checklist' approach was used, 'ticking off' descriptors to allocate band scores
- Some challenges with overlapping of dimensions



## 2. Focus Group

- Problem scripts re-marked
- Minor changes to scale:
  - Descriptor concerning empathy/respect for both parties added to context dimension
  - CEFR/CV vocabulary descriptor for 'less common/idiomatic expressions' removed
- Possibility of a holistic scale?

## Refocusing of the assessment of intralinguistic mediation

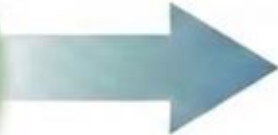


In AoA, the main focus is not language, but rather the fulfillment of relevant tasks in which learners further develop their competencies through the creation of a product embedded in a relevant scenario (North, 2022). The assessment should therefore be task specific and measured in terms of successful task fulfillment.

# Recommendation

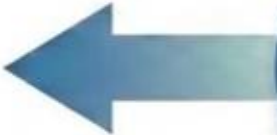


**Rater  
Feedback**






**Holistic Scale  
Approach**

Holistic scale with salient descriptors acting as guidance for raters.



**MFRM  
Results**

# Holistic scale

Traffic light system   

	B1	B2
<b>CONTENT (relevance, accuracy and transformation of ideas to create new meaning).</b>	<p>Main <b>sense</b> of most key issues are addressed (although not necessarily fully developed).</p> <p>Main/Key issues from source texts are <b>relevant</b> and <b>reasonably accurate</b>. However, the focus is on relaying information rather than resolving conflict (<b>solutions may be limited</b>).</p> <p>Ideas are <b>paraphrased</b> in a simple fashion, though some evidence of lifting from input may be present.</p>	<p>All key points are addressed and significant ideas/arguments have been <b>identified</b> and are reliably <b>summarised</b>.</p> <p>Information has been <b>well-paraphrased</b> with little evidence of lifting. There is evidence of additional explanations drawing on prior knowledge (e.g. examples, comparing, contrasting etc.).</p> <p>An obvious attempt has been made to <b>identify areas of common ground/defuse tension</b> and clarify disagreements. <b>Basic solutions or compromises</b> have been suggested. Includes an attempt to facilitate continued communication.</p>
<b>CONTEXTUAL APPROPRIATENESS for communicative purpose (register, layout, style, sociopragmatics).</b>	<p>Register and style are neutral. Layout is simple but clear. Shows <b>some awareness</b> of socio-pragmatics but may include some inconsistencies.</p>	<p>Adapts register, style, and layout effectively to suit communicative purpose and audience most of time. Shows empathy/respect towards both parties. Attempts <b>attenuation</b> and <b>polite language</b> to soften conflict and encourage dialogue.</p>
<b>ORGANISATION (connecting and organising relevant mediated ideas).</b>	<p>Organisation supports overall clarity but may be uneven. Structures ideas with some logical sequencing. Uses a range of basic connectors (e.g. however, because, so).</p>	<p>Organises ideas coherently and cohesively. Uses a variety of connectors and cohesive devices to show relationships between ideas effectively.</p>
<b>COMPREHENSIBILITY (amount of effort and rereading required).</b>	<p>Occasional re-reading may be needed due to awkward phrasing or inaccuracies.</p> <p>Grammar is generally accurate though some errors persist, especially with complex structures.</p> <p>Lexical limitations cause difficulty with formulation at times. Spelling and punctuation are accurate enough to be followed most of the time.</p>	<p>Clear and fluent communication with minimal effort required on behalf of the reader. Text is easy to follow on first reading.</p> <p>Uses some complex sentence forms. Vocabulary range is sufficient to express ideas precisely. Errors do not hinder understanding. Spelling and punctuation are reasonably accurate</p>

# Positive washback

Facilitating the move towards the constructive alignment of planning, teaching and assessment.



Shackleton, C., Turner, N., & Sánchez Cuadrado, A. (2026). Developing a rating scale for written intralinguistic mediation in a local context, *Assessing Writing*, Volume 69, 101049, ISSN 1075-2935, <https://doi.org/10.1016/j.asw.2026.101049>



Developing a rating scale for written intralinguistic mediation in a local context

Caroline Shackleton<sup>a,\*</sup>, Nathan Turner<sup>a</sup>, Adolfo Sánchez Cuadrado<sup>b</sup>

<sup>a</sup> University of Granada, Centro de Lenguas Modernas, Plazuela del Ilustre Vijejo s/n, Granada 18006, Spain  
<sup>b</sup> University of Granada, Faculty of Arts, Carretera Campus, Granada 18071, Spain

# Thank you for listening

# Questions? Comments?