Diseño y validación de escalas: otra evidencia que contribuye a demostrar la validez de nuestras pruebas de dominio

Carmen Álvarez García

mariadelcarmen.alvarez@upm.es

Contenido

- Introducción: principios básicos de la evaluación y de las destrezas productivas

- Protocolo para el diseño de escalas ancladas al MCER

- Validación de las escalas de corrección

Principios básicos de la evaluación

- Fiabilidad
- Validez
  - Apariencia
  - Contenido
  - Respuesta
  - Constructo
- Equidad
- Practicalidad
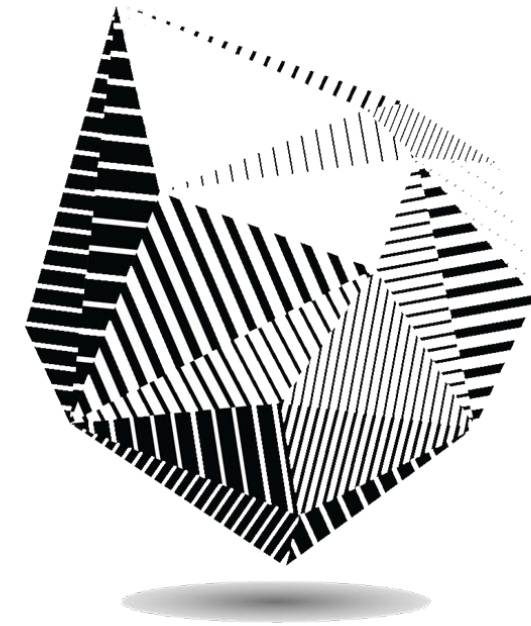
Messick, 1989

Definición de las destrezas productivas

Definición del objetivo del test
- ¿Qué quiero evaluar en la prueba de expresión escrita y expresión oral?

Definición de las habilidades
- Definición del constructo
- Marco teórico – MCER + VC

Toma de decisiones sobre las macro y micro destrezas a evaluar

Operacionalización del constructo

Especificaciones de la prueba y escalas de corrección de las destrezas productivas

Por lo tanto, si nuestro constructo se operacionaliza a través de las escalas de corrección, ¡habrá que comprobar que funcionan!

Evaluación de las destrazas productivas

➢Directa y subjetiva

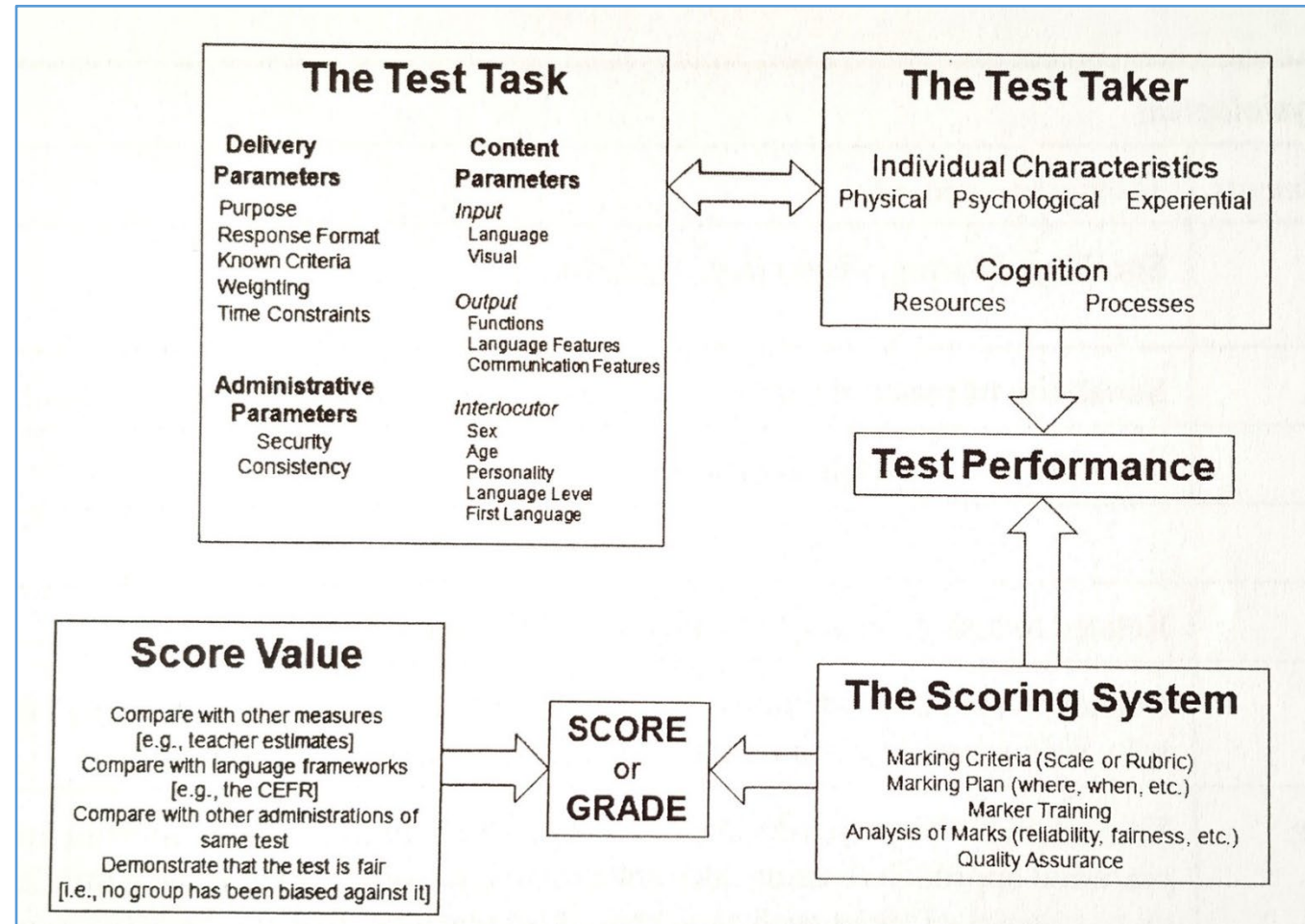- Juicio del corrector sobre la habilidad de los candidatos en función de su rendimiento en una tarea

➢ Herramientas

- Especificaciones
- Escalas de corrección

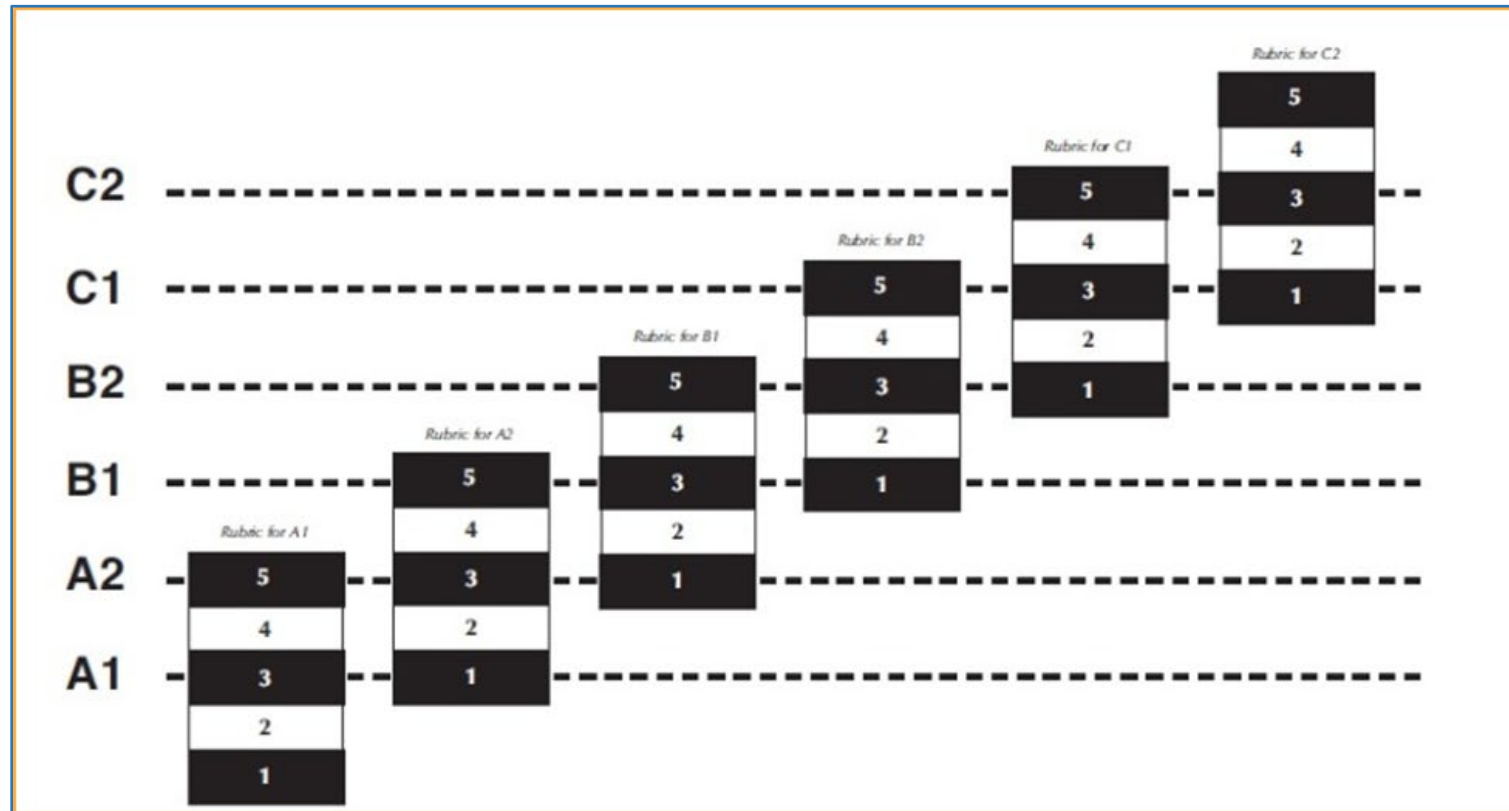➢Variedad de factores involucrados (relación entre ellos)

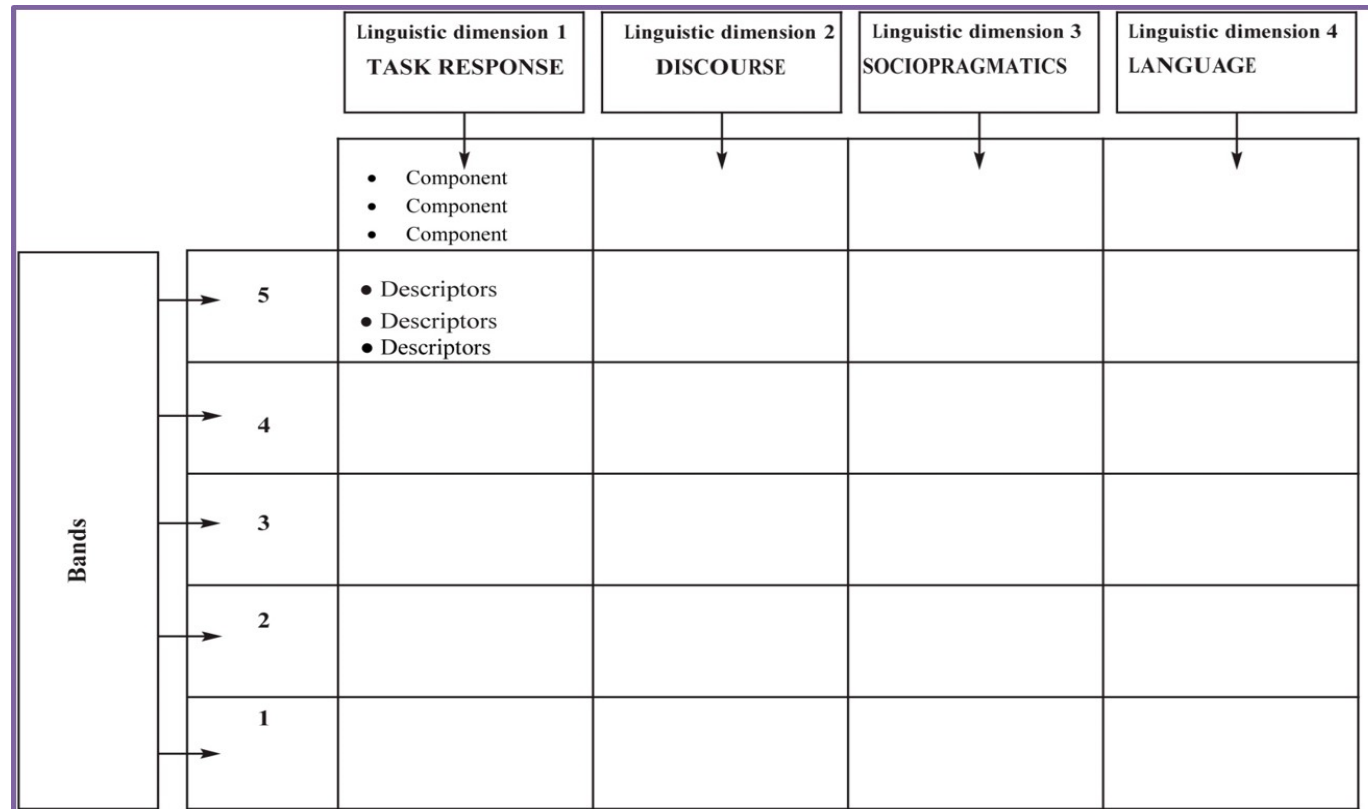➢Papel de los evaluadores

- Formación
- Sesiones de estandarización

O'Sullivan, 2012

Protocolo para el diseño de las destrazas productivas

Fase	Acción
1. Consideraciones iniciales	A. Elegir entre escala holística, analítica o de rasgo primario. B. Identificar el número de dimensiones de la escala y definirlas. C. Establecer el número de bandas. D. Establecer el formato de las calificaciones.
2. Confección de los descriptores	A. Seleccionar las tablas del *MCERL* con información relevante. B. Distribuir los descriptores del *MCERL* como anclaje. C. Confeccionar descriptores de bandas incompletas. D. Revisar el conjunto de los descriptores.
3. Validación 1 (cualitativa)	A. Revisión de expertos B. Modificar la escala en función de lo indicado por los expertos.
4. Validación 2 (cuantitativa)	A. Al menos 2 evaluadores evalúan a 30 candidatos. B. Mediante Rasch se analizan: *data fit, vertical ruler y rating scale category utility*.
5. Implementación	A. Formar a los nuevos evaluadores en el uso de la escala. B. Usar las escalas en condiciones reales de examen. C. Recoger y analizar los datos generados en el examen.
6. Revisión	A. Establecer un ciclo de revisión. B. Recoger los datos generados en diferentes convocatorias. C. Revisar los datos anteriores y repetir la fase 5 si es necesario.

Protocolo para el diseño de las destrazas productivas – consideraciones previas

Cruz, 2024

Protocolo para el diseño de las destrazas productivas – consideraciones previas

	Linguistic dimension 1 TASK RESPONSE	Linguistic dimension 2 DISCOURSE	Linguistic dimension 3 SOCIOPRAGMATICS	Linguistic dimension 4 LANGUAGE
	• Component • Component • Component			
5	• Descriptors • Descriptors • Descriptors			
4				
3				
2				
1				

Bands

Cruz, 2024

Protocolo para el diseño de las destrazas productivas – redacción de descriptores

B1	TASK RESPONSE (CEFR P. 129) Task response/content Elaboration of ideas Information to the reader	DISCOURSE (CEFR P. 61, 125) Thematic development (Text?? And) Information organisation Cohesion	COMMUNICATIVE OBJECTIVE (CEFR P. 122, 124) Register and style Conventions Use of language for an objective	LANGUAGE (CEFR P. 112-114) Vocabulary (range and control) Grammar (range and control) Errors + orthographic control
5	• Content • Simple and some complex ideas on abstract as well as concrete topics are effectively developed. • Detailed information is reliably passed on	• Thematic development shows clear descriptions or narratives, expanding and supporting his/her main points with relevant supporting details and examples. • Information is clearly and coherently organised • A variety of frequent cohesive devices to link his/her utterances into a clear, coherent discourse.	• Formal and informal register are appropriately used. • Conventions are appropriately met/addressed by adjusting the means of expressing to the situation and person(s) concerned. • Language is appropriately used for the situation and there are no crass errors of formulation.	•Vocabulary range and control go from good to high to express most general topics and are sufficient to provide clear descriptions, viewpoints and more abstract or cultural topics such as music and films. • Grammar control is relatively high and displays complex sentence forms. • Errors are minor, not recurring and do not lead to misunderstanding when dealing with simple or complex ideas.
4	• Content • Ideas • Main points in an idea or problem are explained with reasonable precision	• Thematic development is clear in narrative discourse but lacks supporting details and examples. • Information • A limited number of cohesive devices is used to link his/her utterances into a clear discourse occasional imprecisions	• Register • Conventions are addressed by adapting his/her expression to deal with less routine, even difficult situations. • Language is appropriately used for situation and there are no crass errors of formulation.	• Vocabulary range and control are sufficient to express him/herself about topics like family, hobbies and interests, work, travel, current events and abstract cultural topics such as music and films. • Grammar range and control are efficient and accurate in a variety of sentence structures that still display mother tongue influence. • Errors occur, but are very scarce, minor and it is clear what he/she is trying to express, even when expressing more complex and unfamiliar topics
3	• Content • Simple ideas are developed to convey straightforward information • Main points are comprehensibly expressed on a range of familiar subjects within his field of interest,.	• Thematic development is reasonably fluent when relating a straightforward narrative or description. • Information is organised as a connected linear sequence of points • Frequently occurring connectors are used to link a series of shorter, discrete simple elements into a connected, linear sequence of points.	• Neutral register • Most salient textual conventions are appropriately regarded. • Simple language is used flexibly to express much of what he/she and to perform and respond to a wide range of language functions	• Vocabulary range and control are sufficient to express him/herself about topics like family, hobbies and interests, work, travel and current events although lexical limitations may cause repetition. • Grammar range and control are used efficiently throughout simple, correct sentence structures that still display mother tongue influence. • Errors are scarce, minor and do not impede communication when expressing simple thoughts, although major errors still occur when expressing more complex and unfamiliar topics.
2	• Content • Ideas • Reader	• In thematic development he/she has to struggle to relate simple narratives or descriptions. • Information is at some points presented as a linear sequence of points, although with difficulty. • The most frequently occurring connectors are used to link simple sentences, although still not able to narrate a sequence of events.	• Register • Conventions are addressed in a simple but effective way using the simplest common expressions and following basic routines. • Memorised simple phrases are adapted to particular circumstances in order to perform and respond to basic language functions.	• Vocabulary is sufficient to conduct routine, everyday activities involving familiar situations and topics beyond simple survival needs. • Grammar is limited to correct sentence structures that display heavy influence of the mother tongue. • Errors are frequent and intrude with simple and complex ideas, although it is usually clear what he/she is trying to say
1	• Content • Simple ideas are reduced to direct exchanges of limited information. • Main points can only be communicated on familiar and routine matters but in other situations he/she generally has to compromise the message	• Thematic development is reduced to the description of something as a simple list of points. • Information is presented as a simple list of points • Groups of words are linked with simple connectors like 'and', 'but' and 'because'.	• Everyday polite forms of greeting and address. • Difficulty to handle textual conventions for social exchanges. • Simple language is used to perform and respond to limited language functions such as invitations, suggestions and apologies.	• Vocabulary is sufficient for the expression of basic communicative needs and for coping with simple survival needs. • Grammar is limited to simple, correct structures, basic sentence patterns and memorised phrases and brief everyday expressions to satisfy simple needs. • Errors are systematic (tends to mix up tenses and forget to mark agreement). Nevertheless, it is usually clear what he/she is trying to say.

Protocolo para el diseño de las destrazas productivas – redacción de descriptores

B1	TASK RESPONSE (CEFR P. 129) Task response/content Elaboration of ideas Information to the reader	DISCOURSE (CEFR P. 61, 125) Thematic development (Text?? And) Information organisation Cohesion	COMMUNICATIVE OBJECTIVE (CEFR P. 122, 124) Register and style Conventions Use of language for an objective	LANGUAGE (CEFR P. 112-114) Vocabulary (range and control) Grammar (range and control) Errors Orthographic control
5	• Content is relevant to the task and all tasks are satisfactorily addressed. • Simple and some complex ideas on abstract as well as concrete topics are effectively developed. • Detailed information is reliably passed on and the reader is fully informed.	• Thematic development shows clear descriptions or narratives, expanding and supporting his/her main points with relevant supporting details and examples. • Information is clearly and coherently organised, and paragraphs are effectively arranged to avoid unclear connections between ideas. • A variety of frequent cohesive devices to link his/her utterances into a clear, coherent discourse.	• Formal and informal register and style are appropriately used according to the communicative situation. • Conventions are appropriately met/addressed by adjusting the means of expressing to the situation and person(s) concerned. • Language is appropriately used for the situation and there are no crass errors of formulation.	• Vocabulary range and control go from good to high to express most general topics and are sufficient to provide clear descriptions, viewpoints and more abstract or cultural topics such as music and films. • Grammar control is relatively high and displays complex sentence forms. • Errors are minor, not recurring and do not lead to misunderstanding when dealing with simple or complex ideas.
4	• Content is generally relevant to the task and all tasks are addressed although some points might not be fully developed. • Simple and some complex ideas on abstract as well as concrete topics are developed. • Main points in an idea or problem are explained with reasonable precision, so the reader is informed in some detail.	• Thematic development is clear in narrative discourse but lacks supporting details and examples. • Information is coherently organised and clearly arranged in paragraphs although there are some unclear connections between ideas. • A limited number of cohesive devices is used to link his/her utterances into a clear discourse although there might be occasional imprecisions	• Use of formal and informal register and style is generally appropriate although at some points disregarded. • Conventions are addressed by adapting his/her expression to deal with less routine, even difficult situations. • Language is appropriately used with ease for the situation and there are no crass errors of formulation.	• Vocabulary range and control are sufficient to express him/herself about topics like family, hobbies and interests, work, travel, current events and abstract cultural topics such as music and films. • Grammar range and control are efficient and accurate in a variety of sentence structures that still display mother tongue influence. • Errors occur, but are very scarce, minor and it is clear what he/she is trying to express, even when expressing more complex and unfamiliar topics
3	• Content is generally relevant although there might be some minor irrelevances or omissions. All tasks are attempted, but mainly in a general way. • Simple ideas are developed to convey straightforward information • Main points are comprehensibly expressed on a range of familiar subjects within his field of interest, so the reader is on the whole informed.	• Thematic development is reasonably fluent when relating a straightforward narrative or description. • Information is organised as a connected linear sequence of points clearly arranged in paragraphs. • Frequently occurring connectors are used to link a series of shorter, discrete simple elements into a connected, linear sequence of points.	• Neutral register is appropriately used in most cases, although there is an attempt to adapt the style to the communicative situation. • Most salient textual conventions are appropriately regarded. • Simple language is used flexibly to express much of what he/she and to perform and respond to a wide range of language functions effectively.	• Vocabulary range and control are sufficient to express him/herself about topics like family, hobbies and interests, work, travel and current events although lexical limitations may cause repetition. • Grammar range and control are used efficiently throughout simple, correct sentence structures that still display mother tongue influence. • Errors are scarce, minor and do not impede communication when expressing simple thoughts, although major errors still occur when expressing more complex and unfamiliar topics.
2	• Content is generally relevant although there might be some irrelevances or omissions. All tasks are attempted, although some are only mentioned. • Simple ideas are put together to convey straightforward information. • Main points are at least mentioned on a range of familiar subjects within his field of interest, so the reader might need more information to fully understand the message.	• In thematic development he/she has to struggle to relate simple narratives or descriptions. • Information is at some points presented as a linear sequence of points, although with difficulty. • The most frequently occurring connectors are used to link simple sentences, although still not able to narrate a sequence of events.	• Neutral register is appropriately used in most cases, although there is an attempt to adapt the style to the communicative situation which is sometimes unsuccessful. • Conventions are addressed in a simple but effective way using the simplest common expressions and following basic routines. • Memorised simple phrases are adapted to particular circumstances in order to perform and respond to basic language functions.	• Vocabulary is sufficient to conduct routine, everyday activities involving familiar situations and topics beyond simple survival needs. • Grammar is limited to correct sentence structures that display heavy influence of the mother tongue. • Errors are frequent and intrude with simple and complex ideas, although it is usually clear what he/she is trying to say
1	• Content is at some points irrelevant. All tasks are attempted, although some are only mentioned. • Simple ideas are reduced to direct exchanges of limited information. • Main points can only be communicated on familiar and routine matters but in other situations he/she generally has to compromise the message, so the reader is not appropriately informed.	• Thematic development is reduced to the description of something as a simple list of points. • Information is presented as a simple list of points, and somehow arranged in paragraphs although not connected. • Groups of words are linked with simple connectors like 'and', 'but' and 'because'.	• Only aware of the most common everyday polite forms of greeting and address. • Difficulty to handle textual conventions for social exchanges. • Simple language is used to perform and respond to limited language functions such as invitations, suggestions and apologies.	• Vocabulary is sufficient for the expression of basic communicative needs and for coping with simple survival needs. • Grammar is limited to simple, correct structures, basic sentence patterns and memorised phrases and brief everyday expressions to satisfy simple needs. • Errors are systematic (tends to mix up tenses and forget to mark agreement). Nevertheless, it is usually clear what he/she is trying to say.

Protocolo para el diseño de las destrazas productivas – redacción de descriptores

WRITING – B1

Universidad de Jaén | CEALM Centro de Estudios Avanzados en Lenguas Modernas | POLITÉCNICA | CENTRO DE LENGUAS UPM

Band	Language — Vocabulary (range and control) / Grammar (range and control) / Errors	Discourse — Thematic development / Cohesion / Text and information organization	Sociopragmatics — Register / Conventions / Flexibility in the use of language	Task response — Content / Elaboration of ideas / Information to the reader
5	• Vocabulary range and control are good to express general topics and specialist vocabulary of their field, and are sufficient to provide clear descriptions and viewpoints. Is able to vary the formulation to avoid frequent repetition. • Grammar range and control are relatively high. Uses complex sentence forms which may display rigidity or inaccuracy and occasional errors. • Errors are scarce, minor and do not intrude communication when expressing simple or more complex and unfamiliar topics. Spelling and punctuation are reasonably accurate but may show signs of mother tongue influence.	• Thematic development shows clear descriptions or narratives, expanding and supporting his/her main points with relevant supporting details and examples as well as expressing viewpoints and developing clear arguments. • The cohesion devices used are limited in number but help to link utterances into clear, coherent discourse, though there may be some "jumpiness" in a long contribution. • Information is clearly and coherently organised, and paragraphs are clearly and logically arranged to avoid unclear connections between ideas. The difference between fact and opinion is clearly signalled.	• Formal and informal register is appropriately identified according to the communicative situation, but no always so appropriately used. • Conventions are appropriately met by adjusting the means of expressing to the situation and person(s) concerned. • As regards flexibility, language is effectively used for the communicative situation and there are no crass errors of formulation although their language may lack expressive power and idiomaticity.	• Content is relevant to the task and all aspects required are satisfactorily addressed. • Some simple and some complex ideas on abstract as well as concrete topics are effectively developed. • Detailed information is reliably passed on, so that the reader is fully informed.
4	• Vocabulary range and control is sufficient to describe unpredictable situations, explain the main points in an idea or problem with reasonable precision, and express thoughts on abstract or cultural topics such as music and film. • Grammar range and control are efficient and accurate in a variety of less familiar sentence structures that still display mother tongue influence. • Errors are scarce, minor and do not intrude communication when expressing simple ideas. Minor errors still may appear when expressing more complex and unfamiliar topics.	• Thematic development shows clear descriptions and chronological sequences in narratives, and arguments are well developed to be followed without difficulty, but lacks supporting details. • Displays cohesion to introduce a counter-argument in a simple discursive text (e.g. with "however"). • Information is coherently organised and clearly arranged in paragraphs although there are some unclear connections between ideas.	• Use of formal and/or informal register is generally appropriate to the communicative situation although it is occasionally disregarded. • Conventions are addressed by adapting his/her expression to deal with less routine, even difficult situations. • In terms of flexibility, language is used appropriately for the communicative situation and there are no crass errors of formulation.	• Content is generally relevant to the task and all aspects required are addressed although some points might not be fully developed. • Mainly simple ideas on abstract as well as concrete topics are appropriately developed and the expression of some complex ideas is attempted. • Main points in an idea or problem are explained with reasonable precision, so that the reader is informed in some detail.
3	• Vocabulary range and control are sufficient to express him/herself about topics like family, hobbies and interests, work, travel and current events. Lexical limitations may cause repetition. • Grammar range and control are used efficiently throughout simple, correct sentence structures that still display mother tongue influence. • Errors are scarce, minor and do not intrude communication when expressing simple thoughts, although major errors still occur when expressing more complex and unfamiliar topics.	• Thematic development is reasonably fluent when relating a straightforward description or narrative. • Displays cohesion to form longer sentences and links them together using a limited number of cohesive devices. Can link a series of short, discrete simple elements into a connected, linear sequence of points. • Information is organised as a connected linear sequence of points clearly arranged in paragraphs.	• Neutral register is appropriately used in most cases, although there is an attempt to adapt it to the communicative situation. • The conventional structure of the text type is regarded and shows awareness of the most salient conventions. • Simple language is used flexibly to express much of what he/she wants to and to perform and respond to a wide range of language functions effectively.	• Content is generally relevant to the task although there might be some minor irrelevances or omissions. All aspects required are attempted, but mainly in a general way. • All simple ideas are developed to convey straightforward information. • Main points are comprehensibly expressed on a range of familiar subjects within the candidate's fields of interest, so that the reader is informed on the whole.
2	• Vocabulary is sufficient to write about familiar situations and topics beyond simple survival needs. • Grammar range and control are sufficient in simple structures, but mistakes are frequent and there is heavy influence of the mother tongue. • Errors are frequent but do not occur more than once in basic aspects (e.g. tense choice, agreement or word choice). It is usually clear what he/she is trying to say.	• In thematic development he/she has to struggle to relate simple narratives or descriptions as a linear sequence of points, and he/she occasionally succeeds. • Cohesion is achieved through the most frequently occurring connectors, used to link simple sentences in order to tell a story or describe something as a simple list of points. • Information is organised as a connected linear sequence of points and is occasionally fragmented.	• Register is limited to basic language functions (e.g. information exchange, requests, when expressing simple opinions). • Conventions are correctly used to exchange information by text message, by e-mail or in short letters, responding to questions from the other person (e.g. about a new product or activity). • In terms of language flexibility, can adapt memorised, simple phrases to particular circumstances through limited lexical substitution.	• Content is not always relevant to the task and there are noticeable irrelevances or omissions. At least half of the aspects required are attempted, but mainly in a general way. • Several simple ideas are developed to convey straightforward information but does not always succeed at making the point. • Main points are expressed to develop simple ideas which do not always inform the reader.
1	• Vocabulary is sufficient for the expression of basic written communication and for coping with simple survival needs. • Grammar is limited to memorised phrases and brief everyday expressions used to satisfy simple needs. • Errors are systematic in basic aspects (e.g. tends to mix up tenses, forgets to mark agreement or chooses wrong words). Nevertheless, the message is clear for the most part.	• Thematic development is reduced to describing something as a simple list of points. • Cohesion is achieved by linking groups of words with simple connectors like 'and', 'but' and 'because'. • Information is fragmented for the most part.	• Register is noticeably limited to very short social exchanges, using everyday polite forms of greeting and address. • Conventions are correctly used only to convey personal information of a routine nature, for example when introducing themselves. • Language flexibility is hindered by the lack of basic resources.	• Content is not relevant to the task and there are noticeable irrelevances or omissions. Only isolated aspects required in the task are attempted. • Only isolated and simple ideas are developed • The points expressed are scarce and for the most part they do not inform the reader.

Validación cualitativa

1. Revisión de las escalas por parte de expertos

2. Ajuste y modificación de descriptores en base a los juicios de los expertos

Validación cuantitativa

Pilotaje de las escalas con muestras reales ya calibradas

1. Análisis estadístico

   Teoría Clásica de los Test (TCT) (Classic Test Theory)

   Teoría de Respuesta al Ítems (TRI) (Item Response Theory)

2. Ajuste y modificación de descriptores en base a los resultados

Validación cualitativa

EXPRESIÓN ESCRITA B1 — CRITERIOS DE VALORACIÓN — TABLA DE REFERENCIA

B1	10-9-/5 muy competente	8-7 //4	6-5// 3-2.5 competente	4-3 //2	2-1 // 1 no satisfactorio	0 // 0 no asiste no lleva a cabo la tarea respuesta completamente memorizada.
CUMPLIMIENTO DE LA TAREA (* Funciones comunicativas específicas : recomienda, compara, solicita, ejemplifica, etc.) (en las instr)	☐-Alcanza el objetivo con eficacia y claridad. Impresión muy positiva. ☐-Desarrolla todos los puntos a tratar de forma clara, con cierta amplitud y creatividad (incluida extensión). Todo el contenido está presente y es pertinente. ☐-Realiza un claro intento de usar diferentes formulaciones de las funciones que se esperan, a menudo con éxito. ☐-Sigue el tipo de texto y formato sin apenas alteraciones. ☐-Muestra una cierta adecuación de la formalidad del registro.		☐-El objetivo comunicativo se percibe con facilidad. Posibles vacilaciones u omisiones en temas más abstractos. ☐-Aporta información suficiente para desarrollar los puntos más importantes a tratar (incluida la extensión). Contenido por lo general pertinente, posibilidad de algunas inconsistencias. ☐-Expresa la mayoría de las funciones* que se esperan. ☐-Por lo general responde al tipo de texto y formato. Puede haber inconsistencias. ☐-Utiliza un registro en general neutro.		☐-Alcanza una parte del objetivo. Texto compuesto de frases y oraciones muy sencillas y aisladas. El enfoque a menudo distrae al lector. ☐-Desarrolla una parte de los puntos a tratar. Una parte del contenido aportado es pertinente. El texto se aleja de la extensión debida. ☐-Maneja algunas de las funciones que se esperan en su versión más básica. ☐-Una parte del texto responde al tipo y formato que la tarea exige. ☐-No hay descriptor aplicable	☐-El texto no se corresponde con la situación planteada. ☐-No sigue los puntos a tratar. El contenido no es pertinente a la tarea. Extensión tan corta que no hay material calificable. ☐-No es capaz expresar las funciones mínimas que se esperan o estas presentan serias deficiencias. ☐-El texto responde a un tipo y formato completamente diferente. Registro - no hay descriptor aplicable.
ORGANIZACIÓN COHERENCIA Y COHESIÓN mecanismos cohesivos* referenciales: deícticos, relativos, pronombres conectores de inicio, mantenimiento y final de discurso, signos de puntuación.	☐-Produce un texto organizado y claro en párrafos lógicos. ☐-Desarrolla ideas con coherencia y añade información pertinente. Marca la relación entre las ideas con poca rigidez ☐-Hace buen uso de mecanismos de enlace y de cohesión de nivel B1, creando una secuencia lineal eficaz. Utiliza de forma bastante consistente los signos de puntuación.		☐-En general organiza un texto extenso en párrafos sencillos y, en su mayoría, lógicos. ☐-Desarrolla con bastante coherencia ideas sencillas y concretas, algunas de cierta longitud. Distingue entre ideas principales y secundarias. ☐-Uso correcto, en general, de un número limitado de mecanismos de enlace y de cohesión* del nivel en una secuencia cohesionada y lineal. Puntuación con posible influencia de L1 pero no impide la comprensión del texto.		☐-Parte del texto presenta cierto orden. Hay un intento de organización en párrafos. ☐-Algunas ideas muestran un cierto desarrollo. Conecta algunas ideas en una línea coherente. ☐-En ocasiones utiliza los mecanismos de enlace y de cohesión más básicos, hay repetición; en ocasiones conectan y referencian lo que deben. Signos de puntuación con influencia de L1.	☐-El discurso aparece a menudo desordenado. La falta de lógica en los párrafos impide la comprensión del texto. ☐-El texto producido se limita a una serie de frases sencillas y aisladas, apenas hay desarrollo. ☐-Ausencia casi completa de los elementos de enlace y de cohesión. Abundantes errores en la puntuación que impiden la comprensión del texto.
COMPETENCIA GRAMATICAL	☐-Utiliza un repertorio gramatical muy amplio con bastante precisión. Incluye ejemplos de sintaxis más compleja. ☐-Usa con flexibilidad las estructuras gramaticales del nivel e incluso fórmulas más complejas para expresar lo que quiere. ☐-Mantiene un buen grado de corrección de estructuras gramaticales sencillas y algunas más complejas. Todavía hay algún fallo o error no sistemático.		☐-Usa un repertorio gramatical de B1 razonablemente amplio, simple pero en gran medida correcto, para expresar situaciones cotidianas. Texto con influencia de L1. ☐-Usa con cierta flexibilidad las estructuras gramaticales sencillas del nivel. ☐-Mantiene un grado de corrección gramatical suficiente. Posibles errores, incluso alguno sistemático, fallos en la estructura de frase pero por lo general su intención queda clara.		☐-Utiliza un repertorio gramatical sencillo y limitado. A veces consigue transmitir lo que necesita. ☐-Usa parte de las estructuras gramaticales del nivel. Se expresa, por lo general, con poca flexibilidad y repetición. ☐-Usa correctamente una pequeña parte de las formas gramaticales sencillas del nivel. Los errores presentes en frases y oraciones son frecuentes y/o dificultan la comprensión.	☐-Su limitado repertorio gramatical no le permite expresar lo que necesita. ☐-No hay material calificable. ☐-Hay errores gramaticales abundantes y/o sistemáticos de gravedad que con frecuencia impiden la transmisión de las ideas.
COMPETENCIA LÉXICA	☐-Maneja un repertorio léxico amplio sobre temas cotidianos. Usa, a veces con éxito, léxico menos común y algunas expresiones idiomáticas.. ☐-Mantiene un alto control léxico sencillo. Puede cometer alguna imprecisión en temas menos habituales y pensamientos más complejos. ☐-Su ortografía en general correcta, Los errores léxicos y ortográficos son escasos y no impiden la transmisión de la idea.		☐-Utiliza un repertorio léxico razonablemente amplio sobre temas cotidianos y de su interés. El intento de utilizar léxico menos conocido puede presentar limitaciones. ☐-Mantiene un control suficiente de léxico elemental aunque existe cierta repetición y comete errores en temas menos habituales y pensamientos más complejos. ☐-Su ortografía es satisfactoria para el nivel y la tarea. Los errores léxicos y ortográficos no son muy frecuentes y no impiden la comprensión.		☐-Posee un repertorio léxico relativo a situaciones cotidianas que a veces le permite transmitir información básica. ☐-A veces muestra un cierto control del léxico más elemental; existe un alto grado de repetición. ☐-Los errores léxicos y ortográficos con frecuencia provocan malentendidos.	☐-Utiliza un repertorio léxico muy limitado que resulta insuficiente para cumplir con la tarea. ☐-Falta de control sobre léxico elemental y/o interferencia de otras lenguas que dificultan en gran medida la comprensión del mensaje ☐-Comete errores ortográficos graves y sistemáticos que provocan frecuentes malentendidos.

13

Validación cuantitativa - TCT
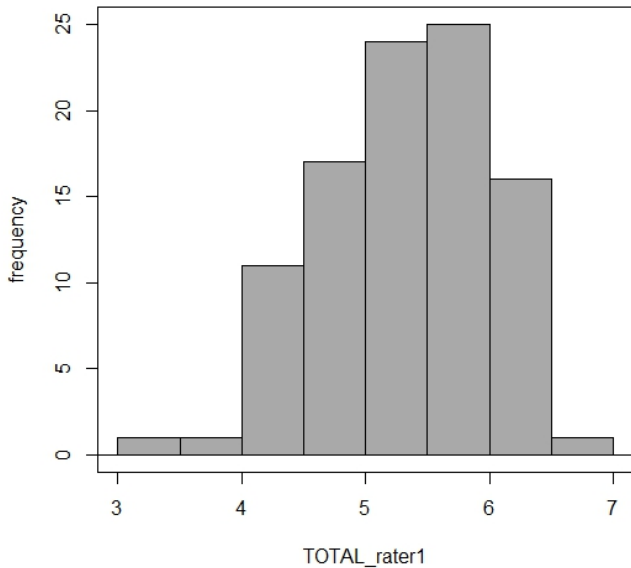
Figure 1. Rater 1 total scores distribution

TOTAL_rater1

Table 1. Rater 1 total scores statistics

		Rater1_Total
N	Valid	96
	Missing	0
Mean		5.46
Median		5.50
Mode		5.50
Std. Deviation		.700
Skewness		-.383
Std. Error of Skewness		.246
Kurtosis		-.114
Std. Error of Kurtosis		.488
Range		4

Table 2. Rater 1 statistics on each criterion

		T.RESPrater1	COH.COHrater1	LEXIC.RESrater1	RANGEACrater1
N	Valid	96	96	96	96
	Missing	0	0	0	0
Mean		5.55	5.34	5.66	5.28
Median		5.50	5.00	6.00	5.00
Mode		5	6	6	6
Std. Deviation		.881	.831	.752	.817
Skewness		-.209	-.273	.057	-.446
Std. Error of Skewness		.246	.246	.246	.246
Kurtosis		.263	-.252	-.402	-.536
Std. Error of Kurtosis		.488	.488	.488	.488
Range		4	4	3	4

Validación cuantitativa - TCT

Table 5. Correlations

Spearman's rho		T.RESP rater2	COH.COH rater2	LEXIC.RES rater2	RANGE AC rater2
T.RESP rater1	Correlation Coef.	.668**	.496**	.493**	.542**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	96	96	96	96
COH.COH rater1	Correlation Coef.	.651**	.700**	.512**	.609**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	96	96	96	96
LEXIC.RES rater1	Correlation Coef.	.536**	.389**	.569**	.568**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	96	96	96	96
RANGE AC rater1	Correlation Coef.	.569**	.465**	.668**	.628**
	Sig. (2-tailed)	.000	.000	.000	.000
	N	96	96	96	96

**. Correlation is significant at the 0.01 level (2-tailed).

Table 6. Correlations

			rater1total	rater2total
Spearman's rho	rater1total	Correlation Coefficient	1.000	.793**
		Sig. (2-tailed)	.	.000
		N	96	96
	rater2total	Correlation Coefficient	.793**	1.000
		Sig. (2-tailed)	.000	.
		N	96	96

**. Correlation is significant at the 0.01 level (2-tailed).
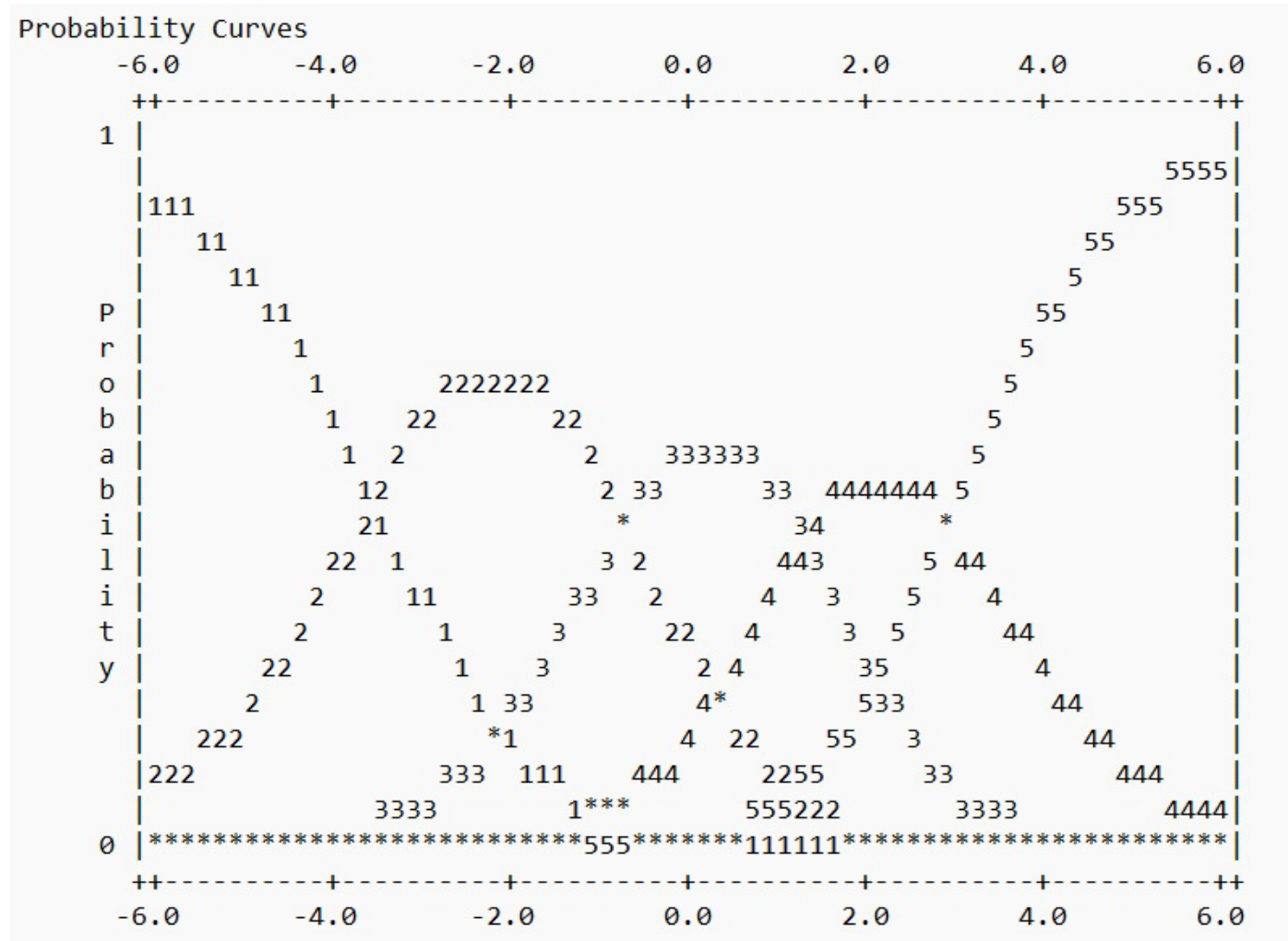
Validación cuantitativa – TRI

Modelo de Rasch – Many-Facet Rasch Measurement (Linacre, 1999 – extensión de uno de los modelos de Rasch)– Programa FACETS (Linacre, 2009)

Puntos críticos en el análisis estadístico de una escala – *Investigating rating scale category utility* (Linacre, 1999)

|	A	BANDA	# DE OBSERVACIONES	RASCH-ANDRICH Threshold	DIFERENCIA	OUTFIT MNSQ
1				WRITE - B1		
3		1	18			0,9
4	Dispersión de dimensiones (en *logits*): 0,27	2	47	-4,86	4,05	1,0
5		3	96	-0,81	2,7	0,8
6		4	101	1,89	1,9	1,0
7		5	98	3,79		1,0
9				WRITE - B2		
10		BANDA	# DE OBSERVACIONES	RASCH-ANDRICH Threshold	Difference	OUTFIT MNSQ
11		1	9			1,0
12	Dispersión de dimensiones (en logits): 1,29	2	56	-3,4	2,6	1,1
13		3	124	-0,8	1,91	1,0
14		4	125	1,11	1,99	1,0
15		5	46	3,1		0,8
17				WRITE - C1		
18		BANDA	# DE OBSERVACIONES	RASCH-ANDRICH Threshold	Difference	OUTFIT MNSQ
19		1	7			1,3
20	Dispersión de dimensiones (en logits): 0,45	2	61	-4,87	4,22	1,1
21		3	138	-0,65	2,51	1,0
22		4	111	1,86	1,79	1,0
23		5	43	3,65		0,8

Validación cuantitativa – TRI

Curvas de probabilidad

Validación cuantitativa – TRI

Vertical ruler
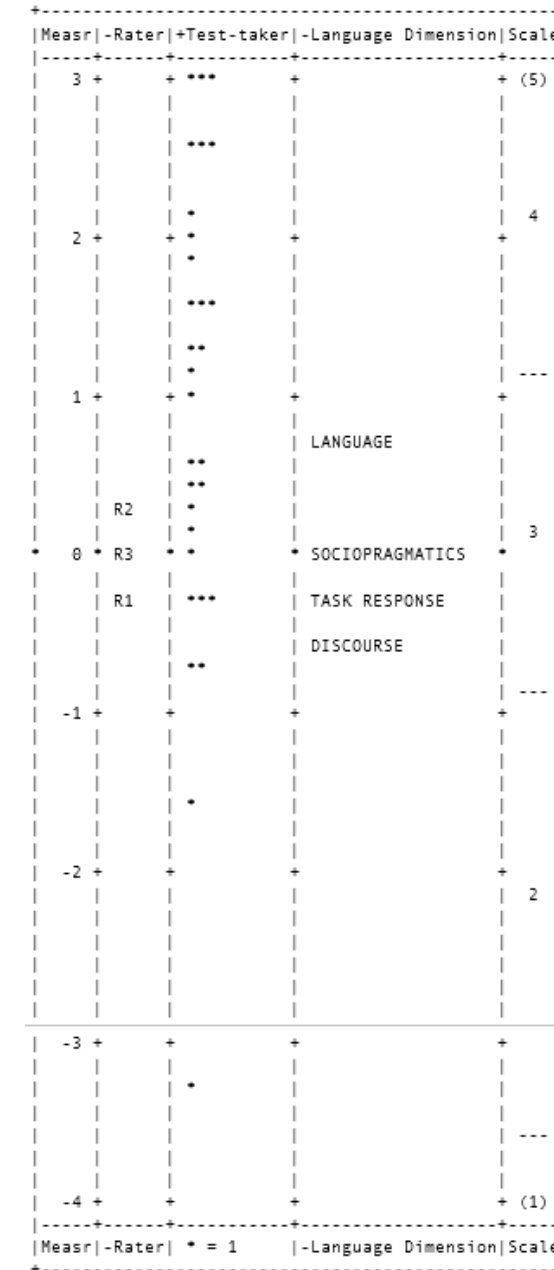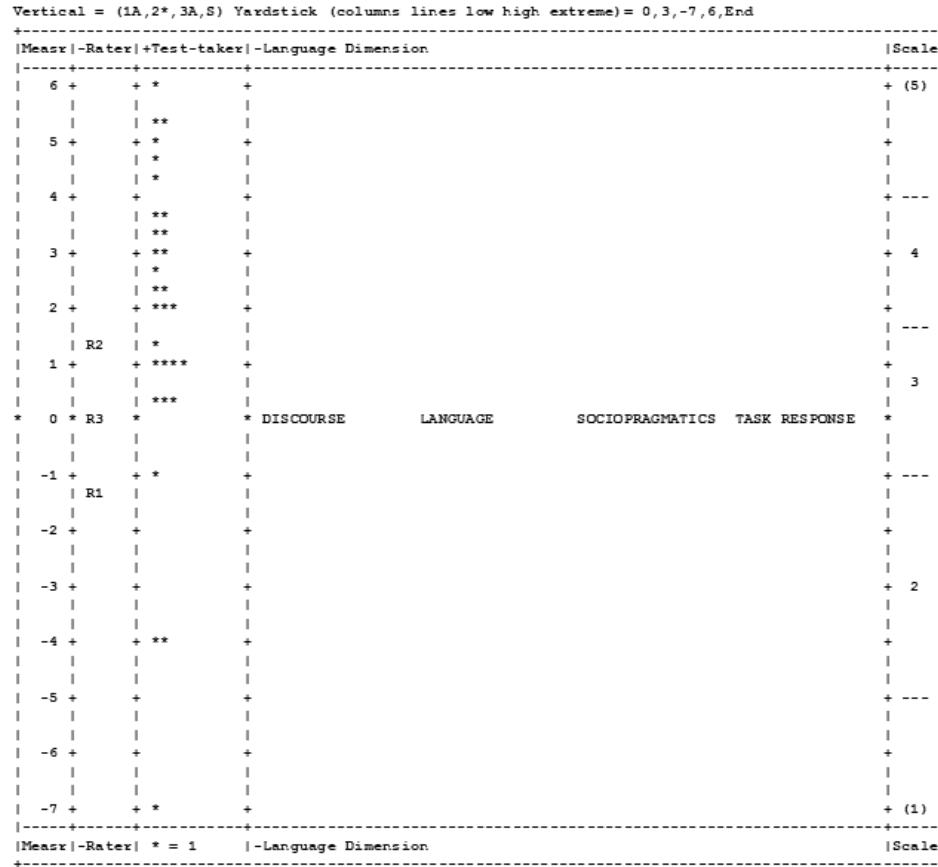
References

Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge, UK: Cambridge University Press.

Bachman, L F. (2004). *Statistical analysis for language assessment*. Cambridge: Cambridge University Press.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge, UK: Cambridge University Press.

Council of Europe. (2020). *CEFR Companion Volume with new descriptors*. Strasbourg, FR: Language Policy Programme.

Cruz, J. (2016). *A protocol to design a CEFR-linked proficiency rating scales for oral production and its App implementation*. Unpublished doctoral dissertation, University of Jaén, Spain.

Cruz Trapero, J (en prensa). *Diseño y validación de exámenes de dominio de lengua*, Madrid: Ministerio de Educación Formación Profesional y Deportes.

East, M. (2009). Evaluating the reliability of a detailed analytic scoring rubric for foreign language writing. *Assessing Writing*, 14(2), 88-115. https://doi.org/10.1016/j.asw.2009.04.001

Eckes T. (2011). *Introduction to Many-Facet Rasch Measurement*. New York: Peter Lang.

References

Green, R. (2013). *Statistical analyses for language test developers*. Basingstoke: Palgrave Macmillan.

Linacre, J. M. (1999). Investigating rating scale category utility. *Journal of Outcome Measurement* 3.2: 103–122

Messick, S. (1989). Validity. In Linn, R. (ed.), *Edutational Measurement*. (pp. 13-103). New York: Macmillan.

Muñiz, J. (2010). Las teorías de los tests: Teoría clásica y teoría de respuesta a los ítems [Test theories: Classical test theory and item response theory]. *Papeles del Psicólogo, 31*(1), 57–66.

O'Sullivan (2012). Assessing Speaking. In C. Coombe, P. Davidson, B. O'Sullivan, & S. Stoynoff (Eds.), *The Cambridge Guide to Second Language Assessment* (pp. 234-246). Cambridge, UK: Cambridge University Press.

Prieto Adánez, G., & Díaz Velasco, A. (2011). Uso del modelo de Rasch para poner en la misma escala las puntuaciones de distintos tests. *Actualidades En Psicología, 19*(106), 5–23. https://doi.org/10.15517/ap.v19i106.43

Prieto Adánez, Gerardo (2011). Evaluación de la ejecución mediante el modelo Many-Facet Rasch Measurement. *Psicothema*, vol. 23, núm. 2, pp. 233-238.

¡Muchas gracias!

Carmen Álvarez García
mariadelcarmen.alvarez@upm.es