



Catálogo de las Buenas Prácticas



A practical guide for constructing rating scales to assess communicative language competences

Eleanor Occeña Gallardo
Universidad Autónoma del Estado de Hidalgo
eoccena@uah.edu.mx

Resumen

El maestro no es un diseñador de exámenes. Hay muchos conceptos en el diseño de exámenes que el estudiante necesita aprender y practicar antes de que pueda diseñar reactivos y construir un criterio para medir el desempeño del estudiante en las actividades de una instrucción basada en competencias. El diseño de exámenes es complejo, y Hughes (2003) tiene la razón en decir que ‘la construcción de una escala de calificación no es un asunto fácil’ (p. 104). Por eso, es necesario proporcionar al estudiante una guía práctica para que pueda medir justamente el desempeño de futuros alumnos en actividades comunicativas.

Palabras clave: Escalas de calificación, constructo, rúbricas, niveles de rendimiento

Abstract

Teachers are not test designers. There are so many concepts to learn and to practice about testing before becoming good at designing test items in performance-based instruction. And these have to be mastered before being able to write down criteria that will measure the student’s performance in communicative language tasks. If designing tests is complex, then Hughes (2003) is right in pointing out that “constructing a valid rating scale is no easy matter” (p.104). Thus, it is necessary to provide a guide that is practical for teachers so they are able to justly measure the student’s communicative performance.

Keywords: Rating scales, construct, rubrics, descriptors, performance levels

Características generales de la Buena Práctica (BP)

Está diseñada para un tipo de curso curricular de modalidad presencial, el objetivo que persigue el curso es la formación de profesores y el nivel de dominio de la lengua al inicio del curso es B2 de acuerdo con el Marco Común Europeo de Referencia para las Lenguas (MCER).

Contexto

1. **Nombre del curso:** Evaluación en la enseñanza y aprendizaje de idiomas
2. **Tipo de curso:** Curricular
3. **Modalidad del curso:** Presencial
4. **Objetivo persigue el curso:** La competencia de diseñar reactivos para exámenes y escalas de calificación
5. **Duración en horas:** 61-80
6. **Distribución del tiempo:** 3 horas los miércoles y 2 horas los viernes; a partir del segundo periodo parcial hasta los finales, es totalmente práctica, en la mayor parte del tiempo en trabajo colaborativo.
7. **Nivel de dominio de la lengua al inicio del curso:** B2

Contexto de la práctica

Número promedio de alumnos: 21-30

Habilidad(es) comunicativa(s), académicas o digitales a las que responde su BP: competencia de diseñar reactivos para exámenes y escalas de calificación.

Necesidades, dificultades y/o particularidades académicas que lo llevaron a decidir que era conveniente implementar la intervención didáctica: Una enseñanza basada en competencias requiere que la instrucción este orientada en actividades comunicativas. Estas actividades requieren la evaluación del desempeño del estudiante con una escala de calificaciones. Lo que es correcto y justo es que el maestro mismo construye cada escala, describiendo el nivel de desempeño deseado basado en las características de la actividad comunicativa. La BP es para guiar al maestro en construir su propia escala de calificación para evaluar competencias comunicativas.

Factores afectivos que incidieron en el diseño/implementación de la BP: Motivación, auto-concepto, autoestima y actitudes y creencias.

Objetivos: Los estudiantes, futuros maestros en la enseñanza de una lengua extranjera, serán capaces de construir sus propias escalas de calificación para evaluar el desempeño de sus estudiantes en las competencias comunicativas con la ayuda de una guía práctica.

Descripción de la BP

Procedimiento

Los estudiantes del curso que imparto están en el séptimo semestre de la carrera y poseen el nivel B2 del Marco Común Europeo de Referencia. Sin embargo, esta guía está diseñada para el maestro de cualquier nivel que está enseñando con el enfoque basado en competencias.

El proceso para la construcción de escalas de calificación se presenta debajo en orden cronológico, seguido por una explicación de cada etapa. En algunas etapas, se presentan ejemplos. Después de la explicación de todas las etapas, encontrarán un ejemplo completo para que puedan visualizar la estructura de la guía.

1. Indicar el *Construct* (la competencia a evaluar)

De acuerdo a De Jong y Harsch (2016), se debe basar un Criterio de Evaluación en el *construct* y en los objetivos de aprendizaje. Mucho antes de leer a De Jong y Harsch, he enfatizado en mi salón de clase que el *construct* es la punta de lanza del diseño de exámenes, así como de la construcción de una escala de evaluación. Cuando el maestro sabe bien lo que quiere evaluar, va a poder trazar el objetivo, describir la operación, escribir la instrucción y luego determinar la escala de calificación. Los *constructs* se expresan como habilidades o competencias. Ejemplo: *can understand/write about/speak...../ is able to*

2. Indicar el Objetivo de la evaluación

Saber lo que se va a evaluar – the *construct* -, permite enunciar el Objetivo – lo que da la justificación de la evaluación. Después de indicar el *construct*, el maestro debe de tener claro el propósito de la evaluación.

3. Describir la *Operación*

La *operación* se refiere a la actividad que el maestro elija para poder evaluar el estudiante. Esta es una importante etapa del proceso: la descripción de la actividad que el estudiante tiene que realizar para hacer operativo el *construct*. Esta tiene un doble propósito: 1) permite que el estudiante demuestre las habilidades y/o las competencias comunicativas de la lengua, para que 2) el maestro pueda evaluar el desempeño de estas habilidades de la lengua.

4. Escribir la *instrucción*

La *operación* es el enunciado de lo que el estudiante tiene que hacer en la evaluación, la *instrucción* es el enunciado dirigido al sustentante de la evaluación. Debe mencionar exactamente los detalles descritos en la *operación* y enunciado en la forma imperativa.

5. Construir la escala de calificación (Marcaje de evaluación)

Hay solo tres pasos a seguir para la *Construcción de la Escala de Calificación*.

1) Tener en cuenta siempre tu *construct*

El *construct* es la guía para la construcción de la Escala de evaluación. A partir de aquí, verán ejemplos.

Ejemplo de un *Construct*: *Can write about a place they have visited and describe what things are interesting of the place*

¿Qué es lo que quiere evaluar de este *construct*?

Ejemplo:

- a. ¿Qué desempeño de competencia/habilidad quiere evaluar/medir? - *the student's ability to write about an interesting place they have visited, describing five things they found interesting.*
- b. ¿Qué aspecto del idioma debe usar el sustentante? - *language to describe places (adjectives), present perfect*

- c. ¿Qué otros detalles mencionaste en la Operación que quiere incluir en la evaluación? -
number of words

2) Determina si la escala de calificación es holística o analítica (nota: La diferencia entre *Holística* y *Analítica* se explicará en los fundamentos en el apartado num.15

Si es holística, vea como se construye las Escalas en el ejemplo en el Anexo 1.

Nota: El número de niveles deseados de desempeño y los descriptores o rúbricas correspondientes dependerán del juicio del maestro y qué niveles deseados de desempeño espera de los estudiantes. El número de puntos serán de acuerdo al número de descriptores con el mayor puntaje para el nivel más alto deseado de desempeño.

Si es analítica: Primero, determinar qué *componentes de habilidades* quiere evaluar? Ejemplo con el mismo *construct* en el apartado de Holístico: *Student is able to write five interesting things about a place they have visited and describe why you found these things interesting.*

Ejemplos de componentes de habilidades:

Componente 1: TASK COMPLETION

Componente 2: USE OF THE LANGUAGE

Segundo, determinar los niveles deseados de desempeño que espera de los estudiantes en cada componente

Vea como se construye las Escalas en el ejemplo en el Anexo 2

Notas:

- En esta escala de calificación para el Componente 1 – Task Completion, hay solo tres descriptores, y los puntos correspondientes son 5-3-1. No se visualiza los números 4 y 2 porque 4 tendrá características del 5 y del 3. Ejemplo: El estudiante podrá escribir sobre un lugar que ha visitado (escala 5), pero excedió en el número de palabras especificado (escala 3), o al revés, el estudiante cumplió con el número de palabras pero describió menos de 5 cosas interesantes del lugar. La escala 2 tendrá las características del 3 y 1
- Las escalas de calificación pueden ser sencillas con pocos descriptores. Lo que se debe recordar es que los niveles deseados de desempeño son los niveles que el maestro espera de sus estudiantes.
- La escala más alta de desempeño no tiene que ser de un nativo del idioma. El maestro decide el nivel más alto de la escala, permitiendo algún margen de error en cualquier aspecto de las habilidades.

Efectos en los estudiantes

Las reflexiones de mis estudiantes:

- Pensaba que los criterios de evaluación se los daban los maestros. No tenía idea que podía construir uno siguiendo la guía de mi maestra.

- Me parece fácil escribir rúbricas con la guía
- Fue fácil escribir rúbricas para la escala analítica
- Al principio me confundían los conceptos, pero en la práctica, la guía lo hizo fácil
- Fue un reto escribir las rúbricas, pero con la guía y la práctica, ya puedo escribir mis propias rúbricas

El concepto del constructo fue totalmente nuevo para mí, y no sabía que con esto me guiaría en el diseño exámenes y las escalas de calificación

Fundamentos teóricos y metodológicos

La descripción de los niveles deseados de desempeño comúnmente se conoce como rúbricas, que se conceptualizan como descripciones narrativas que se separan en niveles de posibles desempeños relacionadas a una actividad (teAchnology, sf).

Estas descripciones de niveles de desempeño conforman el criterio para evaluar el desempeño de los estudiantes en las actividades comunicativas de la lengua. Se han dado diferentes términos en la literatura para referirse al conjunto de descripciones que se usan para evaluar el desempeño del estudiante en actividades comunicativas. Fulcher (en Jong y Harsch, 2016) explica que las escalas son descripciones del criterio de evaluación en “varias niveles ascendentes o bandas, que conforma las bases para juzgar” el desempeño del estudiante. Davies *et al* (1999) afirma que una escala de calificación es una:

...descripción del dominio de la lengua que consiste de una serie de niveles construidos en donde el desempeño del estudiante en una lengua se califica...los niveles o bandas comúnmente son caracterizadas en términos de lo que el sujeto puede hacer con el idioma...y su dominio de las características lingüísticas (p.153)

Hughes (2003, pág. 105) sugiere los siguientes pasos en la construcción de una escala de calificación para cualquiera de las habilidades productivas:

- ¿Cuál es el propósito de la evaluación?
- ¿Cuántas diferentes habilidades se evaluarán?
- ¿Cómo se reportarán los resultados?
- ¿Cuáles van a ser los componentes de la habilidad que quiere evaluar?
- El criterio de evaluación es analítico*¹ o holístico**², o ambos

Materiales y recursos

Principalmente una computadora para diseñar la guía y un proyector de datos para explicar la guía

¹ El criterio Analítico es dividir el juicio en varios sub-habilidades (Baxter, 1997). Con una escala analítica, los que califican tienen que tomar decisiones en varios componentes de desempeño separadamente en lugar de una sola calificación para el desempeño total. (Alderson *et al.* 1995; Arter & McTighe, 2001; Weigle in Chinda 2013).

² El criterio Holístico es la asignación de una sola calificación a una habilidad, por ejemplo, a una escritura o una expresión oral, con base a la percepción del desempeño en su totalidad (Hughes, 2003. pág. 94).

Características de su BP

En el segundo periodo parcial es cuando mis estudiantes empiezan a poner en práctica la guía para la construcción de escalas de calificación, al mismo tiempo que aprenden a diseñar exámenes. Antes de la segunda evaluación parcial, los estudiantes hicieron dos prácticas. En la primera práctica ellos realizaron la construcción de una escala holística y una escala analítica para un mismo examen. En la segunda práctica, ellos eligieron la escala que mejor evaluará el desempeño del estudiante basando en el constructo. Para la segunda evaluación parcial, ellos diseñaron dos exámenes – uno para una habilidad receptiva y otro para una habilidad productiva. En esta evaluación, tuvieron que elegir y construir la escala de calificación más adecuada para la evaluación de cada habilidad. Su tarea para la evaluación final fue diseñar exámenes para las cuatro habilidades con sus correspondientes escalas de calificación. En todas estas prácticas, el estudiante construye su escala siguiendo la guía y tuvieron retroalimentación en cada trabajo.

Esto es la estrategia de enseñanza y aprendizaje que he llevado en los últimos 5 semestres. Considero que esta propuesta es una BP porque una guía para la construcción de escalas de calificación le enseña a uno como personalizar su propia escala de calificación y no depender en escalas elaboradas por otros o copiadas del internet. Muchas veces las rubricas son genéricas y se prestan a interpretaciones diversas. He aprendido como docente que además de enseñar los conceptos del curso, hay que demostrar la operatividad de estos.

Referencias

- Baxter, A. (1997). *Evaluating your Students*. London: Richmond Publishing.
- Chinda, B. (2013). Considerations in Performance-Based Language Assessment: Rating Scales and Rater Training. *PASAA*, 46, 141-148.
- Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T., y McNamara, T. (1999). *Dictionary of language testing* (Vol. 7). Cambridge: Cambridge University Press.
- De Jong, N. y Harsch, C. (2016). Rating the Constructs – Assessment Criteria Rating Scales. *EALTA Summer School*, Innsbruck. Recuperado de http://www.ealta.eu.org/events/Summer_school_2016/06_EALTA%20SuSch%202016_Rating%20criteria%20scales.pdf
- Hughes, A. (2003). *Testing for Language teachers*. Cambridge University Press.
- teAchnology (sf). *Why Rubrics?* Recuperado de <http://www.teach-nology.com/tutorials/teaching/rubrics/>

Anexos: Evidencias o materiales de apoyo

1. Ejemplo de una escala de calificación Holística
2. Ejemplo de una escala de calificación Analítica
3. Estructura cómo seguir la Guía

<http://catalogo-buenas-practicas.portal.anui.es.mx/wp-content/blogs.dir/71/files/sites/71/2019/08/Anexos-A-practical-guide-for-constructing.pdf>