

ESTUDIO SOBRE LA APLICACIÓN DE DIFERENTES INSTRUMENTOS DE EVALUACIÓN DE LA ESCRITURA ACADÉMICA EN LENGUA EXTRANJERA (INGLÉS)

Romano, M. Elisa; Martínez, Julia I.

Facultad de Lenguas, UNC

romano.mariaelisa@gmail.com, juliamartinez@gmail.com

INTRODUCCIÓN

La evaluación de la escritura en lengua extranjera constituye un campo de acción complejo, que suele generar ansiedad y aprehensión entre docentes debido a los efectos e implicancias tanto a nivel pedagógico como ético y político. Es por esto que el impacto de la evaluación y los modos de evaluar en las prácticas áulicas es una problemática susceptible de ser analizada y explorada en profundidad. Los resultados de un proyecto de investigación realizado en 2014-2015 sobre principios y creencias que guían la aplicación de criterios de evaluación de ensayos por parte de los docentes de *Lengua Inglesa II* evidenciaron diferencias no sólo en la naturaleza de los criterios de evaluación utilizados por los docentes, sino también en las prioridades que establecen en el proceso de evaluación de las producciones textuales de los estudiantes. El objetivo de este trabajo es presentar los primeros avances del proyecto para 2016-2017, en el que nos proponemos diseñar y analizar la aplicación de dos instrumentos diferentes de evaluación de ensayos académicos en lengua extranjera para determinar su confiabilidad y utilidad.

MARCO TEÓRICO Y ANTECEDENTES

La evaluación de la escritura objeto de esta investigación (evaluación de ensayos en exámenes) puede considerarse, siguiendo a Goodman & Swann (2003) y Hyland (2003), como una *evaluación sumativa*, dado que se califica formalmente y se lleva a cabo, generalmente, al final del proceso de enseñanza con el fin de determinar si se han logrado los objetivos planteados. Este tipo de evaluación es identificada, además, como una *evaluación directa* de la escritura; es decir, se evalúa la escritura a través una producción textual individual y no a través de ejercicios de componentes lingüísticos que hacen a la escritura, lo cual se define como *evaluación indirecta* de la escritura (Williams, 2005).

Por otra parte, dos términos de uso frecuente en el área de evaluación y fundamentales para el estudio que nos proponemos son confiabilidad (*reliability*) y validez (*validity*). El primero de estos conceptos hace referencia a la medida en la cual una determinada respuesta (un texto, en este caso) será evaluada siempre de la misma manera con los mismos resultados (Ferris & Hedgecock, 2014). Una tarea de escritura puede considerarse confiable si mide lo que intenta medir de manera consistente tanto en lo que respecta a los estudiantes (el mismo estudiante en diferentes ocasiones) como en lo que respecta a los evaluadores (la misma tarea a la vista de distintos evaluadores) (Hyland, 2003). El segundo concepto implica, a grandes rasgos, que la prueba debe evaluar lo que se pretende evaluar, y no otras habilidades y/o contenidos (Williams, 2005).

En lo que respecta a abordajes de evaluación de un texto, una distinción básica es la que existe entre el *enfoque holístico* y el *enfoque analítico*. La *evaluación holística* aborda al texto como un todo y conlleva la asignación de una nota o evaluación final general. Esta calificación se basa en la premisa de que el constructo de la escritura es una entidad que puede capturarse mediante una escala única que integra las cualidades inherentes de la escritura y que la calidad sólo puede ser reconocida por lectores expertos (y no mediante instrumentos objetivos) (Hamp-Lyons & Kroll, 1996; Hamp-Lyons, 2003). En cambio, el *enfoque analítico* implica la identificación de distintos aspectos del texto, de modo de determinar las fortalezas y debilidades de cada uno de los aspectos, para luego establecer una evaluación final. Este enfoque implica la calificación de rasgos múltiples y considera el constructo de la escritura como un fenómeno complejo y multifacético. Una de las ventajas de esta forma de evaluar los textos es que le permite a los evaluados analizar las fortalezas y debilidades en sus propios textos, y de esta manera, orientar estrategias de revisión y mejora de los propios productos y procesos de escritura (Deane et al., 2008; Hamp-Lyons, 2003; Williams, 2005).

Algunos estudios recientes en evaluación de la escritura parecen comprobar uno de los principios expresados por Perkins (1983): la evaluación holística es más efectiva o útil cuando el objeto de evaluación es la capacidad o habilidad general del estudiante para escribir (en lengua extranjera o materna), más que una producción textual en particular (Panou, 2013; Rezaei & Lovorn, 2010). Más allá de estos resultados, e independientemente del enfoque que se favorezca, parece existir un acuerdo generalizado sobre la importancia de

asegurar la consistencia y equidad en la aplicación de criterios por parte de los evaluadores. Teniendo en cuenta estas investigaciones y considerando la naturaleza particular de nuestro contexto educativo, nos propusimos determinar qué instrumentos de evaluación demuestran ser más confiables en instancias de evaluación sumativa en *Lengua Inglesa II*.

METODOLOGÍA

En esta investigación participaron todos los docentes de la cátedra *Lengua Inglesa II* de las carreras de Profesorado, Traductorado y Licenciatura de la Facultad de Lenguas, UNC, que están a cargo de la evaluación de ensayos en instancias de evaluación: ocho docentes en total.

Para el cumplimiento de los objetivos planteados se propuso una metodología que combina aspectos cuantitativos y cualitativos, a fin de determinar qué instrumentos de evaluación demuestran ser más efectivos en nuestro contexto particular de escritura en lengua extranjera.

Por un lado, durante 2016 se diseñaron y pilotearon dos instrumentos de evaluación (una escala holística y una escala analítica), para lo que se tuvieron en cuenta estudios anteriores e investigaciones previas (Ferris & Hedgecock, 2014; Kern, 2000; Mei, 2010; Shaw & Weir, 2007), así como también los resultados obtenidos en el proyecto de investigación 2014-2015 sobre criterios docentes y ponderación en la evaluación de la escritura. Por otro lado, se programó la aplicación de los instrumentos de evaluación de la siguiente manera y en dos momentos diferentes.

En una primera instancia, en abril de 2017, se seleccionaron 10 ensayos escritos por estudiantes en la segunda evaluación escrita del año lectivo 2016. Dicha selección fue realizada por una evaluadora experta. Durante el mes de mayo, se llevó a cabo una sesión de entrenamiento en la que se acordó la implementación de la escala holística para la evaluación de los 10 ensayos seleccionados. La escala diseñada consta de cinco bandas: las bandas 3, 4 y 5 describen textos que cumplen, en distintos grados, con los objetivos de la tarea y pueden considerarse aprobados, y las bandas 1 y 2 describen ensayos que no logran los objetivos de la tarea y que están, por lo tanto, aplazados.

Cada docente evaluó dichos textos de manera individual en el término de una semana y luego entregó los textos evaluados. Una vez recolectados los resultados de las evaluaciones

de todos los docentes, se registraron los datos en una planilla de Excel y, actualmente, se están analizando estadísticamente a fin de establecer el grado de correspondencia o acuerdo entre los evaluadores.

En una segunda instancia, está prevista la repetición del procedimiento, pero aplicando una escala analítica para la evaluación de la escritura, para fines de julio de 2017. Al finalizar esta segunda instancia de evaluación, se llevará a cabo un encuentro grupal con todas las docentes participantes a fin de recolectar opiniones, actitudes, apreciaciones y/o valoraciones en relación con la utilidad de cada uno de los instrumentos utilizados.

Finalmente, en una tercera etapa, los resultados obtenidos en ambas instancias serán comparados y contrastados y evaluados en términos de consistencia entre evaluadores.

RESULTADOS PRELIMINARES

Los datos correspondientes a la aplicación de la escala holística están siendo procesados, pero en la Tabla 1 podemos observar algunos resultados preliminares:

	E 1	E 2	E 3	E 4	E 5	E 6	E 7	E 8	E 9	E 10
Banda 5										1
Banda 4	1					1		1	1	2
Banda 3	6	1	1	1		4	1	4	5	5
Banda 2	1	7	5	5	7	2	5	3	2	
Banda 1			2	2	1	1	2			

Tabla 1: Calificación de ensayos utilizando una escala holística

Esta tabla contiene parte de los resultados obtenidos, específicamente el número de evaluadores que seleccionó cada una de las bandas de la escala holística aplicada para calificar cada uno de los 10 ensayos. En color se indica el mayor número de evaluadores que coincidieron en la selección de cada una de las bandas. Como podemos observar, en ningún caso encontramos coincidencia entre los 8 evaluadores; el número mayor es 7, en el caso de los ensayos 2 y 5, y la correspondencia se da en la banda 2, es decir, en el caso de un ensayo en el que la mayoría de los evaluadores consideró que el texto estaba desaprobado.

En lo que respecta a las bandas que describen ensayos aprobados, la mayor coincidencia se observa en el caso del ensayo 1, al cual 6 de los 8 evaluadores calificaron

de acuerdo a la banda 3; otros casos de coincidencia se evidencian en los ensayos 9 y 10, en los que 5 evaluadores coincidieron en aplicar la banda 3. De estos dos últimos casos el que nos indica mayor uniformidad entre evaluadores es el correspondiente al ensayo 10, ya que los tres evaluadores que no aplicaron la banda 3, aplicaron las dos bandas superiores respectivamente, es decir, en este caso todos los evaluadores consideraron que el ensayo estaba aprobado.

Los ensayos 10 y 5 se destacan como los únicos dos casos en los que hay acuerdo entre todos los evaluadores en lo respecta a la aprobación o aplazo de la tarea, ya que, aunque no calificaron con las mismas bandas, todas indican que el ensayo estaría aprobado (E 10) o desaprobado (E 5).

CONCLUSIONES

Aunque todavía faltan las pruebas estadísticas y los datos correspondientes a la escala analítica, los resultados preliminares nos indican que, a pesar de indicar ciertos puntos de acuerdo, hay que continuar trabajando para lograr mayores niveles de confiabilidad entre evaluadores.

Especialmente, parecería que es necesario realizar un trabajo conjunto para establecer los criterios de aprobación y aplazo, a fin de que esa calificación tan significativa y de alto impacto para los estudiantes no presente variaciones, como hemos encontrado en la mayoría de los textos evaluados en esta etapa de nuestra investigación.

Este tipo de trabajos nos permite, además de indagar sobre modos e instrumentos de evaluación más confiables, reflexionar sobre las propias prácticas evaluativas, lo que constituye el primer paso hacia la incorporación de nuevas pautas.

REFERENCIAS BIBLIOGRÁFICAS

- Deane, Paul; Odendahl, Norah; Quinlan, Thomas; Fowles, Mary; Welsh, Cyndi & Bivens-Tatum, Jennifer (2008): *Cognitive models of writing: Writing proficiency as a complex integrated skill*. Princeton, NJ: ETS.
- Ferris, Dana & Hedgecock, John (2014): *Teaching L2 Composition. Purpose, process and practice*. Nueva York: Routledge.

- Goodman, Sharon & Swann, Joan (2003): "Planning the assessment of student writing". En: Coffin, Caroline; Curry, Mary Jane; Goodman, Sharon; Hewings, Ann; Lillis, Theresa M. & Swann, Joan (eds.): *Teaching academic writing. A toolkit for higher education*. Estados Unidos: Routledge, pp. 73-100.
- Hamp-Lyons, Liz (2003): "Writing teachers as assessors of writing". En: Kroll, Barbara (ed.): *Exploring the dynamics of second language writing*. Reino Unido: Cambridge University Press, pp. 162-189.
- Hamp-Lyons, Liz & Kroll, Barbara (1996): "Issues in ESL writing assessment: An overview". En: *College ESL*, 6 (1), pp. 52-72.
- Hyland, Ken (2003): *Second language writing*. Nueva York: Cambridge University Press.
- Kern, Richard (2000): *Literacy and language teaching*. Oxford: OUP.
- Mei, Wu Siew (2010): "Investigating raters' use of analytic descriptors in assessing writing". En: *Reflections on English Language Teaching*, 9 (2), pp. 69-104.
- Panou, Despoina (2013): "L2 writing assessment in the Greek School of Foreign Languages". En: *Journal of Language Teaching and Research*, 4, pp. 649-654.
- Perkins, Kyle (1983): "On the use of composition scoring techniques, objective measures, and objective tests to evaluate ESL writing ability". En: *TESOL Quarterly* 17, pp. 651-671.
- Rezaei, Ali Reza & Lovorn, Michael (2010): "Reliability and validity of rubrics for assessment through writing". En: *Assessing Writing*, 15, pp. 18-39.
- Shaw, Stuart & Weir, Cyril (2007): *Examining writing. Research and practice in assessing second language writing*. Cambridge: CUP.
- Williams, Jessica (2005): *Teaching writing in second and foreign language classrooms*. Chicago: McGraw-Hill.