

YU MAN (于漫)

*Universidad de Estudios Internacionales de Shanghai*

---

## EEE-4 como herramienta de efecto colateral: ¿es beneficioso o dañino?<sup>①</sup>

**Resumen:** En este artículo se hace un estudio empírico del Examen del Español como Especialidad-Nivel 4 (EEE-4) organizado por el Comité Nacional de Español del Ministerio de Educación de China. Como una prueba estatal de competencias genéricas en español desarrollada por las universidades chinas, EEE-4 ha tenido 15 años de historia desde 1999. Con el creciente número de participantes, la mayor influencia y el adicional impacto social, su función de evaluar la docencia universitaria la hace una de las pruebas cada vez más importantes en el dominio de español. ¿Cómo se puede promover efectos beneficiosos de ella? ¿Cómo se logra evitar consecuencias negativas o mitigar su impacto dañino? ¿Si existen máximas para alcanzar o monitorizar su impacto? En vista de la doble función de una prueba estandarizada, el análisis se basa en un marco teórico holístico y tiene en cuenta factores y conceptos clave como: el concepto y los modelos de “efecto colateral”, la validación de evaluación y el papel que pueden jugar los diseñadores de prueba. Si los educadores y los proveedores de examen podemos recopilar los datos estadísticos según sea necesario, estudiar y analizar todos los resultados relevantes con motivo de explorar sus efectos colaterales positivos, sería sumamente provechoso para el aprendizaje y la enseñanza, así como el establecimiento de un sistema de evaluación de castellano en China.

**Palabras clave:** análisis empírico, EEE-4 (1999–2013), efectos colaterales

### 1. Introducción

Cuando deseamos conocer cómo transcurren los procesos de aprendizaje en cualquier área curricular es necesario evaluarlos. En este sentido, la evaluación se convierte en una actividad educativa que consiste en recoger datos, valorarlos y tomar decisiones para intervenir en el currículo y cambiarlo en la dirección que se considere oportuna.

Por otra parte, desde hace tiempo se ha afirmado, en una amplia gama de contextos, que las pruebas ejercen una influencia potente sobre los principiantes que se disponen a tomar estos exámenes y en los profesores que tratan de ayudarles a prepararlos. Las siguientes afirmaciones son las típicas reclamaciones:

---

① Este trabajo ha sido financiado por la Fundación de Investigaciones Científicas de la Universidad de Estudios Internacionales de Shanghai (上海外国语大学区域国别研究中心课题“全球化背景下西班牙语语言教育战略研究”的阶段性成果).

There is a natural tendency for both teachers and students to tailor their classes to the demands of the test, especially when it is important to the future of the students and pass rates are used as a measure of the teacher success.

(Buck 1988: 17)

It is common to claim the existence of efectos colaterales (the impact of a test on teaching) and to declare that tests can be powerful determiners, both positively and negatively, of what happens in classrooms.

(Wall & Alderson 1993: 41)

En este trabajo quisiera centrar en las reflexiones sobre los trabajos que llevamos 15 años realizando para hacer un estudio empírico del Examen del Español como Especialidad-Nivel 4. Es una prueba estatal, diseñada, desarrollada y organizada por el Comité Nacional de la enseñanza de Español del Ministerio de Educación de China con motivo de evaluar las competencias genéricas en español de los alumnos universitarios chinos después de dos años de estudio. A partir de 1999, EEE-4 ha tenido 15 años de historia hasta la fecha. El número de participantes del examen va aumentándose, su influencia se hace cada año mayor y su adicional impacto social sigue creciendo. Su función de evaluar la docencia universitaria la hace una de las pruebas cada vez más importantes. La situación nos llevó a indagar sobre la relación entre la prueba oficial y los procesos de enseñanza — aprendizaje del español como lengua foránea. ¿Cómo se puede promover efectos beneficiosos de ella? ¿Cómo se logra evitar consecuencias negativas? O ¿cómo se mitiga su impacto dañino? ¿Si existen máximas para alcanzar o monitorizar su impacto?

En vista de la dualidad de una prueba estandarizada, la investigación se basa en un marco teórico holístico y tiene en cuenta factores y conceptos clave como: el concepto y los modelos de “efecto colateral”, la validación de prueba y el papel que pueden jugar los diseñadores de prueba. Si los educadores y los proveedores de examen podemos recopilar los datos estadísticos según sea necesario, valorar los resultados relevantes con motivo de explorar sus efectos colaterales positivos, sería sumamente provechoso para el aprendizaje y la enseñanza. Sería favorable para el establecimiento de un sistema de evaluación de español en China.

Para desarrollar el contenido vamos a centrar en las recientes investigaciones sobre el impacto de una prueba de lengua extranjera, primero compilando definiciones de efectos colaterales y construcciones relacionadas. A continuación, usaremos un modelo propuesto por Hughes (1993) e impreso en Bailey (1996) como un trampolín para hablar de los participantes, los procesos y los productos implicados en el efecto colateral. Finalmente, comentaremos sobre métodos de investigación potenciales para investigar fenómenos asociados con efectos colaterales y sugeriremos la investigación que se podría hacer, usando EEE-4 como ejemplo de una prueba en gran escala - cuya aplicación proporciona una oportunidad ideal para investigar el posible efecto colateral. A manera de cierre, presentaremos unas conclusiones y recomendaciones derivadas del análisis, teniendo en cuenta el interrogante en el título.

## 2. Efecto colateral e impacto: definiciones

En este apartado se describen los modelos y conceptos antes mencionados, que servirán de fundamento para el análisis de EEE-4 que se hará posteriormente. Iniciamos con el modelo propuesto por Cohen (1994), ya que a la luz de este modelo, nos interesa examinar cuáles son los propósitos generales y cuáles son las razones específicas que fundamentan y justifican la aplicación de dicho examen.



Figura 1: El Propósito de la evaluación: Adaptado de Cohen (1994: 23)

El efecto colateral o secundarios, podemos definirlo, en términos generales, como las consecuencias que tiene la evaluación en los procesos y en los productos del aprendizaje y de la enseñanza. Tales efectos pueden repercutir en las actitudes de los directivos, los docentes, los educandos y los padres de familia entre otros. También pueden tener repercusión en los procesos de enseñanza y aprendizaje, en el currículo, los contenidos, la metodología y en los productos que se derivan de estos procesos.

Algunos autores han analizado estos efectos, han mostrado cómo funcionan, han destacado los efectos positivos y negativos. Aún otros han llegado a considerarlos como uno de los principios más importantes de la evaluación. Según Bachman y Palmer (1996) estos efectos actúan en dos niveles: un nivel micro, cuando el efecto recae directamente en los estudiantes y los docentes; y un nivel macro, cuando el efecto actúa sobre la sociedad y el sistema educativo. En el primer caso, los efectos ya sean positivos o negativos se manifiestan en los procesos de enseñanza y de aprendizaje. De manera particular, éstos pueden generar cambios en los contenidos, en las metodologías y en las formas de aprendizaje o pueden causar una reducción en el currículo. En el segundo caso, los efectos pueden incidir en las políticas educativas a nivel nacional, cuando a partir de los resultados de las evaluaciones se puede controlar y velar por la calidad de la educación, definir los perfiles de los futuros profesionales y asignar recursos económicos según los niveles de desempeño para

el caso de las universidades públicas. A nivel social, los efectos pueden crear expectativas en la comunidad, en la familia, en los empleadores, entre otros.

De acuerdo con Messick (1996:241-256), los efectos colaterales de la evaluación pueden ser o positivos o negativos. Él describe estos efectos como “las pruebas que miden exclusivamente la competencia gramatical, el léxico, la sintaxis, las expresiones de referencia, entre otros, pueden producir efectos negativos en los profesores y en los estudiantes que interactúan en clases con una orientación comunicativa”. En efecto, esta situación trae como consecuencia, al menos, un debilitamiento y una reducción tanto del currículo como de los contenidos, por cuanto el énfasis que se le da a los aspectos formales de la lengua restringe de manera desfavorable su uso en situaciones reales de comunicación. En el caso contrario, cuando la evaluación indaga sobre la competencia comunicativa, teniendo en cuenta de manera integrada sus dos componentes: la competencia organizativa y la competencia pragmática, los efectos que se producen son positivos, pues dependiendo de los resultados de tal evaluación, el profesor y los estudiantes pueden saber si el currículo, los contenidos y los objetivos se ajustan a las orientaciones comunicativas.

Brown (2004) por su parte, incluye el efecto colateral entre los cinco principios de la evaluación según las siguientes definiciones:

*Viabilidad:* se refiere a los medios que se usan en la evaluación y a las circunstancias que la hacen posible; por ejemplo, los aspectos administrativos y económicos, los procedimientos usados y los tiempos asignados para la misma.

*Validez:* es el punto de correlación entre lo que la evaluación mide en la realidad y lo que pretendía o intentaba medir.

*Confiabilidad:* se entiende como la correspondencia entre la evaluación y los resultados obtenidos por los evaluados.

*Autenticidad:* se define como la consistencia que existe entre el éxito o el fracaso del trabajo asignado en la evaluación al momento de ponerlo en práctica o replicarlo en el mundo real como tal.

*El efecto colateral.* Este autor considera que este principio es el más importante y lo concibe como la influencia que tiene la evaluación en la enseñanza y en el aprendizaje.

Para apreciar mejor la importancia atribuida por Brown al efecto colateral y sus múltiples influencias, presentamos en la siguiente figura un modelo que describe de manera integrada las repercusiones de una evaluación sobre los participantes, los procesos y los productos de los procesos de enseñanza y de aprendizaje. Este modelo nos permitirá, más adelante, analizar en el caso que nos ocupa el alcance que tiene el examen EEE-4.

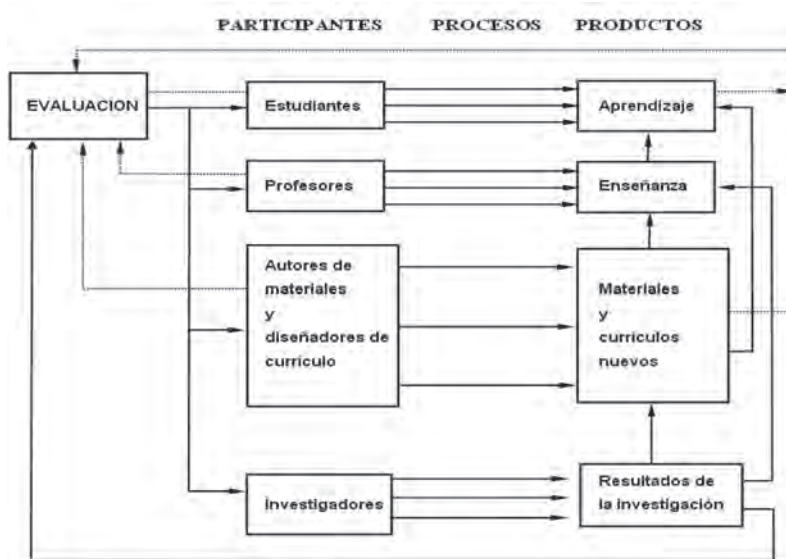


Figura 2: Modelo básico de las 3Ps de efectos colaterales (reimpreso de Bailey, 1996a, p. 264)

Este modelo debe su nombre a la letra inicial de cada una de sus categorías principales, en su versión original en inglés, que coincide con los nombres en español: participantes, procesos y productos. El modelo facilita la comprensión de los efectos que un fenómeno tan complejo como es la evaluación produce en la enseñanza, el aprendizaje y la investigación.

De acuerdo con el autor de esta propuesta, Bailey (1996) describe cada una de las categorías y la forma como éstas se engranan. La categoría *participantes* incluye a los estudiantes y profesores de lenguas, administradores, diseñadores de materiales y editores, de manera general “a todas las personas cuyas percepciones y actitudes hacia su trabajo pueden ser afectadas por la evaluación”. El *proceso* cubre cualquier acción emprendida por los participantes que puede contribuir al desarrollo del aprendizaje. De acuerdo con Hughes (1993), tales procesos incluyen la elaboración de materiales, el diseño, cambios en los métodos o contenidos de la enseñanza, las estrategias de aprendizaje y las técnicas para resolver los exámenes. El *producto* se refiere a aquello que se aprende y a la calidad del aprendizaje.

La naturaleza de la evaluación puede afectar las percepciones y actitudes de los participantes hacia su enseñanza y las tareas de aprendizaje. Estas percepciones, a su vez, pueden afectar aquello que los participantes hacen cuando realizan los procesos, incluyendo las preguntas que se van a encontrar en la evaluación; y estos procesos afectarán los resultados del aprendizaje, el producto de ese trabajo.

### 3. La competencia y modelo de múltiples perspectivas

Por otra parte, como se aprecia en la Figura 2, sobre la competencia del lenguaje Bachman (1990) la divide en dos grandes componentes: la competencia organizadora y la competencia

pragmática. La primera consta de dos subcomponentes: la competencia gramatical y la competencia textual. La competencia gramatical incluye el léxico, la morfología, la sintaxis y la fonología, mientras la competencia textual contiene la cohesión, las referencias, la elipsis, la conjunción, la cohesión léxica u organización retórica.

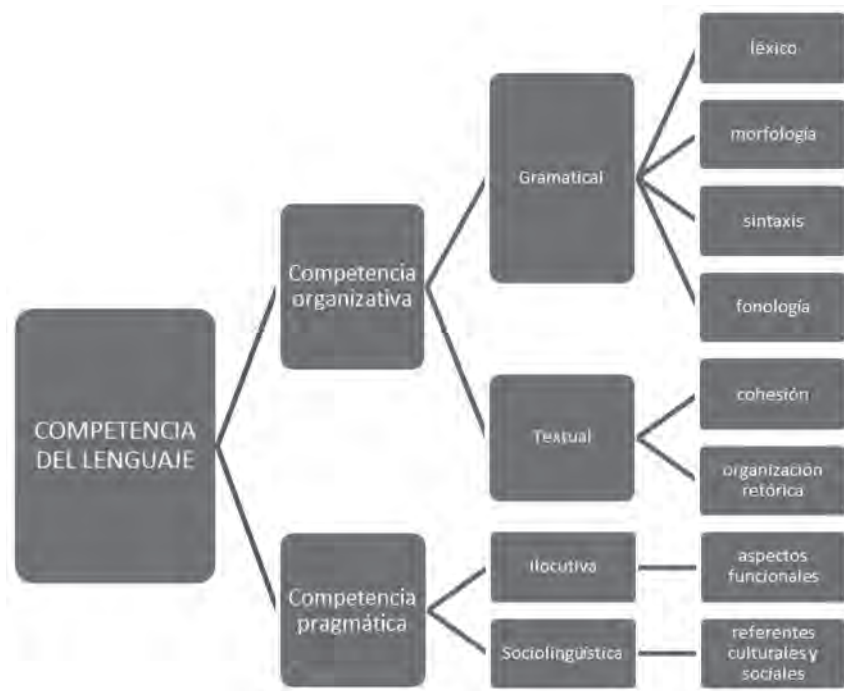


Figura 3: Competencia del lenguaje

La segunda, competencia pragmática, está relacionada con el uso funcional que hacemos de la lengua en situaciones reales de comunicación. A su vez, esta competencia se subdivide en competencia ilocutiva y competencia sociolingüística. La primera abarca aspectos funcionales del lenguaje tales como expresar ideas o sentimientos, entre otros, y la segunda abarca aspectos relacionados con los referentes culturales y sociales, los dialectos y registros, para mencionar algunos de ellos.

Además de las dos competencias generales ya mencionadas, Bachman también hace alusión a una tercera competencia general, la estratégica, que actúa en el uso de la lengua como un componente externo a la competencia del lenguaje. Para este autor la competencia estratégica, más que una competencia del lenguaje, es una habilidad general, que permite a un individuo hacer el uso más efectivo de las habilidades disponibles al llevar a cabo una tarea determinada. Dicha competencia se define como la capacidad de servirse de recursos verbales y no verbales con el objeto tanto de favorecer la efectividad en la comunicación como de compensar fallos que puedan producirse en ella, derivados de lagunas en el conocimiento que se tiene de la lengua o bien de otras



condiciones que limitan la comunicación.

El modelo propuesto por Bachman es un desarrollo de los modelos de Canale & Swain (1980, 1983) y en ambos modelos se ha conservado su carácter integral que nos permite de manera simultánea una aproximación en conjunto a la competencia del lenguaje y a partir de esta competencia hacer una valoración de EEE-4, en particular.

Los anteriores aportes de Bachman, si bien giran alrededor de la competencia del lenguaje en general, presentan algunas coincidencias con el trabajo desarrollado por Bhatia (2002), más focalizado desde la perspectiva del análisis del universo del discurso y del género aplicado, desde múltiples perspectivas. El artículo “Análisis de género aplicado: modelo de múltiples perspectivas” de Bhatia (2002), nos permite llevar a cabo un análisis sobre los tipos de texto o géneros usados en el examen EEE-4 y su pertinencia en las comunidades profesionales. En este trabajo el autor primero describe las propiedades características del “universo del discurso” desde cuatro perspectivas diferentes, pero interrelacionadas: la perspectiva del mundo real, la perspectiva socio-cognitiva del escritor, la perspectiva del analista del discurso y la perspectiva pedagógica. Luego, para moderar las tensiones o traslapos que se presentan entre las perspectivas mencionadas, Bhatia propone un modelo de análisis de género desde tres perspectivas: el discurso como texto, el discurso como género y el discurso como práctica social.

Para Bhatia, las cuatro perspectivas, que describen el complejo, dinámico y variado universo del discurso, producen tensiones y problemas que se traslapan e interactúan. Por ejemplo, la construcción social e individual del discurso presenta una tensión entre las formas discursivas socialmente aceptadas y su manipulación en favor de los intereses personales; el mundo real y el pedagógico nos plantea el siguiente interrogante: ¿Hasta qué punto las prácticas pedagógicas deberían reflejar o dar cuenta de las realidades del mundo del discurso? y el mundo real y analítico, este otro interrogante: ¿Hasta qué punto los análisis deberían dar cuenta de todas las realidades del mundo del discurso?

Las tres perspectivas anteriores sobre el discurso se deben interpretar como elementos que guardan una relación de complementariedad y reciprocidad. Además de ser un modelo integrador, se constituye en uno de los fundamentos de ELE.

La Tabla 1 sintetiza y muestra los alcances de estas tres perspectivas del análisis de género, a través de una pregunta clave, que para cada caso se plantearían desde la perspectiva de la evaluación, en el caso particular del presente trabajo.

|                      |                        |  |  |
|----------------------|------------------------|--|--|
| <b>Discurso como</b> | <b>Práctica Social</b> | <i>¿Cómo las prácticas sociales, las identidades y las estructuras sociales constriñen el discurso?</i>            | <b>Contexto:</b> <i>configurado de manera amplia en términos de las realidades socioculturales</i>           |
|                      | <b>Género</b>          | <i>¿Por qué las personas elaboran el discurso de la forma en que lo hacen y qué hace posible esta elaboración?</i> | <b>Contexto:</b> <i>configurado de manera limitada en términos de los vínculos textuales.</i>                |
|                      | <b>Texto</b>           | <i>¿Qué rasgos del léxico y la gramática se pueden caracterizar estadística y/o funcionalmente?</i>                | <b>Contexto:</b> <i>configurado de manera más específica en términos de las culturas de las disciplinas.</i> |

**Tabla 1: El discurso: múltiples perspectivas. Adaptada de Bhatia V. K. (2002)**

#### 4. Investigando el impacto: el enfoque de EEE-4

EEE-4 es una prueba oficial diseñada, desarrollada y organizada por el Comité Nacional de la Enseñanza de Español del Ministerio de Educación de China con motivo de evaluar las competencias genéricas en español de los alumnos universitarios tras dos años de estudio. En primer lugar, examinaremos, apoyados en el modelo de Cohen, los objetivos de la prueba. Luego, a partir de los dos últimos modelos descritos: “La competencia del lenguaje” propuesta por Bachman y “El modelo de múltiples perspectivas”, sugerido por Bhatia, por su carácter integrador y por la virtud de mostrar de una manera comprensiva la competencia del lenguaje y del discurso; haremos un análisis de qué y cómo se evalúa. Es preciso señalar que los dos modelos en los que se basa el análisis tienen un aspecto fundamental en común: conciben la lengua no sólo como un sistema cerrado, aislado e independiente que tiene una estructura y unos componentes, sino también como un sistema que sirve para expresar funciones comunicativas en el ámbito social, cultural y académico en situaciones reales de interacción. Después de esto, con los resultados del análisis presentaremos, a manera de conclusión, los efectos colaterales que EEE-4 puede producir en los participantes, en los procesos de la enseñanza y del aprendizaje de la lengua extranjera y en los resultados de estos procesos. Finalmente, las recomendaciones formuladas como cierre del artículo están orientadas a producir efectos positivos.

De acuerdo con las características de la enseñanza universitaria de español en China, en el documento oficial de EEE-4, define que dicho examen es “un instrumento estandarizado para la evaluación externa de la calidad de la Educación Superior”. Se hace alusión a la responsabilidad de cumplir funciones de inspección y vigilancia y, además, proporcionar información para el mejoramiento de la calidad de la educación. Para probar con eficacia la capacidad genérica de los estudiantes de usar todas las capacidades básicas, EEE-4 utiliza tareas tanto opcionales como fijas, con ejercicios de rasgo subjetivo principalmente, mientras que los de carácter objetivo son complementarios. Sus tareas están confinadas a la competencia gramatical, la textual y la pragmática del lenguaje.



En el Guía de Orientación de EEE-4 también establece los siguientes objetivos:

Comprobar el grado de desarrollo de las competencias de los estudiantes próximos a culminar dos años de estudio que ofrecen las instituciones de educación superior.

Producir indicadores de valor agregado de la educación superior en relación con el nivel de competencias de quienes ingresan a este nivel; proporcionar información para la comparación entre programas, instituciones y metodologías, y mostrar su evolución en el tiempo.

Servir de fuente de información para la construcción de indicadores de evaluación de la calidad de los programas e instituciones de educación superior, que fomenten la cualificación de los procesos institucionales y soporten el proceso de toma de decisiones en todos los órdenes y componentes del sistema educativo.

Según Cohen (1994), la evaluación cuenta con tres propósitos generales: el administrativo, el educativo y el investigativo. El examen EEE-4 se centra principalmente en el propósito educativo y administrativo y esto se puede constatar en los aspectos específicos como el control y la certificación. Los canales de comunicación de la gama nacional después del examen entre los resultados y los participantes (directivos, docentes, investigadores y estudiantes) permiten una retroalimentación de los procesos de enseñanza y de aprendizaje de la lengua. En cuanto al propósito investigativo, ésta es una de las mejores oportunidades para conocer e indagar sobre los procesos de enseñanza y de aprendizaje de español en nuestro contexto social, cultural y académico y como consecuencia una articulación y fortalecimiento de la relación entre docencia e investigación. Por tanto, puede producir un efecto que aporta beneficios al sistema educativo. Estos propósitos de la evaluación se articulan, se interrelacionan y se cohesionan, pueden lograr ideas innovadoras, transformaciones reales y propuestas de cambio tanto en los marcos conceptuales como en las metodologías usadas para la enseñanza de español.

Dadas estas condiciones, se plantea en el documento de referencia, “evaluar la competencia de español de los estudiantes para comunicarse efectivamente”. Es posible “clasificar a los evaluados” de una manera confiable y de conformidad.

Aquí mencionamos solamente uno de los resultados más interesantes que hemos sacado: el Alfa de Cronbach de EEE-4 en los últimos 14 años. En psicometría el Alfa de Cronbach es un coeficiente que sirve para medir la fiabilidad de una escala de medida, y cuya denominación Alfa fue realizada por Cronbach en 1951. Cuando un investigador trata de medir una cualidad no directamente observable, el Alfa de Cronbach permite cuantificar el nivel de fiabilidad. Cuanto más se aproxime a su valor máximo “1”, mayor es la fiabilidad de la escala. Además, en determinados contextos se considera que valores del alfa superiores a 0,7 o 0,8 son suficientes para garantizar la fiabilidad de la escala. En la tabla siguiente se observa la fiabilidad de EEE-4 de los 14 años pasados:

| Año  | Alpha de Cronbach | N. de tareas | N.de casos |
|------|-------------------|--------------|------------|
| 1999 | .854              | 14           | 124        |
| 2000 | .879              | 15           | 210        |
| 2001 | .850              | 15           | 244        |
| 2002 | .920              | 14           | 312        |
| 2003 | .876              | 13           | 352        |
| 2004 | .892              | 12           | 473        |
| 2005 | .899              | 13           | 448        |
| 2006 | .906              | 14           | 395        |
| 2007 | .913              | 13           | 739        |
| 2008 | .927              | 13           | 1079       |
| 2009 | .914              | 13           | 1259       |
| 2010 | .913              | 13           | 1705       |
| 2011 | .928              | 12           | 2025       |
| 2012 | .920              | 13           | 2615       |

**Tabla 2: Alpha de Cronbach de EEE-4 (1999–2012)**

Al revisar tanto los textos como las preguntas de la prueba, notamos que corresponden a situaciones auténticas de uso que los participantes van a encontrar en el mundo real y logren una alineación real con los descriptores del examen.

Además, a la luz de la perspectiva de Bhatia del mundo real, EEE-4 se caracteriza por su dinamismo y fluidez, propósito que evidentemente se cumple en EEE-4.

## 5. Conclusiones y recomendaciones

A partir de los fundamentos teóricos planteados y el análisis realizado a la prueba de EEE-4, se pueden derivar las siguientes conclusiones y recomendaciones:

El propósito de la prueba de competencias genéricas es predominantemente educativo y administrativo, puesto que cumple funciones tales como la de inspección y la de vigilancia; además, la presentación de la prueba EEE-4 se exige como requisito de grado y ofrece incentivos a los estudiantes que ocupen los primeros lugares en la prueba, entre otros. El propósito investigativo debe ser fortalecido. La prueba EEE-4, de manera concluyente, aparece alineada con las habilidades lingüísticas, y con los descriptores, y con los tipos de texto. Y finalmente aparece alineada con los avances teóricos de la lingüística del análisis del discurso y de género.

De manera categórica, EEE-4 produce un efecto positivo en las creencias y actitudes de los participantes en los procesos de enseñanza y de aprendizaje por las siguientes razones: Esta prueba tiene una concepción de la competencia comunicativa y de la comprensión de lectura, por cuanto la naturaleza de la lengua y de la competencia comunicativa, de una parte, y la complejidad de la comprensión de lectura, de otra, aparecen conceptualizadas de manera holística. Así mismo,

la concepción del currículo, la metodología, los objetivos y los contenidos del área de lengua extranjera no quedan determinados completamente por los alcances de la prueba. Por ejemplo, en el caso de las habilidades lingüísticas, éstas se circunscriben a la comprensión de lectura, aunque en un nivel muy básico por cuanto los tipos de texto y el tipo de actividades corresponden a la perspectiva del mundo real del discurso sugerida por Bhatia.

Finalmente, cualquier acción intencionada producirá consecuencias no deseadas o efectos secundarios. Los efectos colaterales pueden ser de modo positivo o negativo. El objetivo inicial de estas reflexiones era conocer el estatus, los fundamentos teóricos, los objetivos, los contenidos y el tipo de preguntas usados en EEE-4. A través de las investigaciones fueron surgiendo distintas reacciones que, de manera consciente o inconsciente, remitían a la relación existente entre la evaluación oficial y los procesos del aprendizaje y enseñanza de la lengua española, en las universidades de China. A la luz del análisis previo y con el propósito de promover los efectos beneficiosos de ella, y lograr evitar consecuencias negativas o mitigar el impacto dañino de una prueba estandarizada, una de las recomendaciones que podrían ser consideradas fue, tener en cuenta EEE-4 en su estado actual, como un referente de la docencia, pero no transformar los cursos en clases cuyo objetivo principal fuese la preparación de los estudiantes para presentar dicho examen; Es decir, NO enseñar para “aprobar el examen”.

## Bibliografía

- Alderson, C. & Curtis A. (2004): *Washback in Language Testing: Research Contents and Methods*. Mahwah, Lawrence Erlbaum
- Alderson, C. & Wall D. (1993): “Does washback exist?” *Applied Linguistics* (14): 120–121
- Alderson, C. Clapham C. & Wall D. (2000): *Language Testing Construction and Evaluation*, Beijing, Foreign Language Teaching and Research Press
- Bachman, L. & Palmer A. (1996): *Language Testing in Practice*, Oxford, Oxford University Press
- Bachman, L. (1990a): *Fundamental Considerations in Language Testing*, Oxford, Oxford University Press
- Bachman, L. (1990b): “Habilidad lingüística comunicativa. *Competencia comunicativa*,” *Documentos básicos en la enseñanza de lenguas extranjeras*, Madrid, Edelsa
- Bailey, K. (1996): *Working for Washback: A review of the Washback concept in language testing*, *Language Testing*
- Bhatia, V. K. (2002): *Applied genre analysis: a multi-perspective model* (Consultado en enero 2009 en [www.aelfe.org/documents/text4-Bhatia](http://www.aelfe.org/documents/text4-Bhatia).)
- Brown, G & Yule (1983): *Teaching the Spoken Language Acquisition*, Cambridge, Cambridge University Press
- Brown, H. Douglas (2004): *Language Assessment Principles and Classroom Practices*, NY, Pearson Education
- Buck, G. (1988): “Testing listening comprehension in Japanese university entrance examinations,” *JALT Journal* (10): 15–42
- Canale, M. (1983): “On some dimensions of language proficiency,” *Issues in Language Testing Research*, Rowley, Ed. Oller J. W.
- Canale, M. & Swain M. (1980): “Theoretical bases of communicative approaches to second language teaching and testing,” *Applied Linguistics* (1): 1–47
- Centro Virtual Cervantes Instituto Cervantes. Competencia comunicativa. Consultado en octubre de 2009 en (<http://>)

[cvc.cervantes.es/ensenanza/biblioteca\\_ele/diccio\\_ele/diccionario](http://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/diccionario))

Cohen, A. D. (1994): *Assessing Language Ability in the Classroom*, Boston, Heinle & Heinle

Hughes, A. (1989): *Testing for Language Teachers*, Cambridge University Press

Hughes, A. (1993): *Backwash and TOEFL 2000*, University of Reading

Messick S. (1996): "Validity and Washback in language testing," *Language Testing*, (13-3): 241-256

丁黎明(2009): "主、客观测试对英语专业教学反拨效应解析", 《外语学刊》(1)

韩宝成(2002): "语言测试理论、实践与发展", 《外语教学与研究》(4)

黄大勇、杨炳钧(2002): "语言测试反拨效应研究概述", 《外语教学与研究》(4)

孔文、李清华(2009): "TEM4阅读能力构成成分研究", 《现代外语》(3): 287-295

李筱菊(1997): 《语言测试科学与艺术》, 长沙, 湖南教育出版社

刘润清(1992): 《语言测试和它的方法》, 北京, 外语教学与研究出版社

全国高等学校西班牙语专业基础阶段教学大纲编写组(1998): 《高等学校西班牙语专业基础阶段教学大纲》, 上海, 上海外语教育出版社

全国高等学校西班牙语专业四级考试大纲编写小组(2011): 《全国高等学校西班牙语专业四级考试大纲》, 上海, 上海外语教育出版社

杨惠中(1999): 《语言测试与语言教学》, 外语界(1): 16-25

邹申(1998): 《英语语言测试》, 上海, 上海外语教育出版社