

La selección de las tareas de escritura en los exámenes de lengua extranjera destinados al ámbito académico

The selection of writing tasks in second language tests for academic purposes

Arturo Mendoza Ramos

Centro de Enseñanza de Lenguas Extranjeras, Universidad Nacional Autónoma de México .

a.mendoza@cele.unam.mx

Mendoza Ramos, A. (2015). La selección de las tareas de escritura en los exámenes de lengua extranjera destinados al ámbito académico. *Revista Nebrija de Lingüística Aplicada* (2015) 18.

RESUMEN

El objetivo del presente artículo consiste en identificar la pertinencia de las tareas de escritura seleccionadas para un examen de español con fines académicos y también en identificar aquellas variables intrínsecas a la tarea seleccionada que favorecen o no la interacción que se da con el examinado. El estudio se desarrolló en dos fases: pilotaje y primera aplicación del examen. En total participaron 59 estudiantes no hispanohablantes de grado y de posgrado. Los resultados demuestran que las tareas discretas propuestas permiten que los estudiantes empleen las habilidades de escritura indispensables para comunicar el conocimiento escrito dentro del contexto universitario, mas no aquellas requeridas para construirlo. Asimismo, se evidencia que la correcta selección de gráficas y temas para argumentar incide positivamente en la forma en la que el sustentante interactúa con la tarea. Finalmente, cabe resaltar que el modo de transcripción influencia la percepción del evaluador sobre el contenido del texto.

Palabras clave: tareas de escritura, exámenes de lengua con fines académicos, escritura académica

ABSTRACT

The aim of this paper is to identify the relevance of the tasks chosen for a Spanish language exam for academic purposes, and also to identify the task's intrinsic variables that might enhance or not the interaction with the examinee. The study was conducted in two phases: pilot and first administration of the exam. 59 non Spanish speakers at the graduate and undergraduate level took part in the project. The results show that the discrete tasks proposed allow students to use the writing abilities that are important to communicate in the academic context, but not those abilities to build up knowledge. By the same token, the appropriate selection of graphs and argumentative topics promotes the student's interaction and engagement with the task. Finally, it is important to highlight that the transcription mode influences the rater's perception about the content of the script.

Keywords: writing tasks, language exams for academic purposes, academic writing

Fecha de recepción: 10 enero de 2015

Fecha de aceptación: 10 de marzo de 2015

1. INTRODUCCIÓN

Las tareas¹ de escritura que frecuentemente se emplean en los exámenes de lengua para el ámbito académico (ELAA) han sido duramente criticadas por "carecer de autenticidad respecto del dominio para el cual pretenden evaluar". Según los detractores de dicha modalidad (Leki, Cumming y Silva 2008; Moore y Morton 2005; Uysal 2010), ningún estudiante universitario escribe un texto en 30 minutos, a mano y sin conocimiento previo para construir el conocimiento. No obstante, para poder diseñar un examen, se debe describir el universo de generalización a partir del dominio para el cual deseamos evaluar (Kane 2006). En el caso de las pruebas escritas de los ELAA, dicha labor consiste en determinar qué tareas de escritura empleadas para evaluar serían generalizables del universo para el cual deseamos evaluar. Sin embargo, en el caso

del ámbito académico, una de las grandes dificultades reside en lo amplias y diversas que son las prácticas de escritura a nivel universitario (Mendoza en prensa). De hecho, Bachman (2002) menciona que si bien es cierto que las tareas de escritura propuestas para un examen deben estar relacionadas con tareas pertenecientes al mundo real, estas no son reales, aunque sí auténticas.

Lo anterior genera la disyuntiva respecto de establecer un balance entre autenticidad² y estandarización (Bachman 2002; Kane, Crooks y Cohen 1999; Weigle 2002). Estos autores aseveran que en dominios que son muy amplios, la incorporación de tareas con alto grado de autenticidad dificulta tanto su corrección como su estandarización, generándose sesgos en los resultados obtenidos y entorpeciendo el proceso de operacionalización de la misma. Los autores comentan que en los exámenes de alto impacto es importante que estos se encuentren estandarizados, siempre y cuando lo anterior no vaya en detrimento de la autenticidad. Sin embargo, la gran polémica que se ha desatado en contra de los exámenes estandarizados versa justamente en la generalización y toma de decisiones que se efectúan a través de un micro universo seleccionado del dominio en cuestión (Huot, O'Neill y Moore 2010).

Weigle, por ejemplo, menciona cuatro razones por las cuales simple y llanamente las tareas de escritura de ELAA difieren con respecto a un escrito desarrollado en el ámbito universitario:

A typical writing test – in which examinees must write a timed impromptu essay on a topic about which they had no advance knowledge – is relatively inauthentic in at least four ways. First, academic writing outside of a testing situation typically involves the use of source materials as input... Second, except in the case of timed examinations, most academic writing is not timed or speeded, and writers can take as much time... and revise their writing before turning it in for a grade. Third, while writing tests are frequently scored by raters who are unknown to the examinees (and vice versa), the audience for most academic writing is the students' instructor... Finally, most academic writing... is judged primarily on accuracy of content, rather than on the appropriateness of the organization or the use of language (Weigle 2002: 52).

No obstante, pese a que las tareas propuestas para un examen no sean tareas reales que un estudiante enfrentará al momento de cursar estudios universitarios, eso no significa que carezcan de autenticidad. Como bien afirma Bachman (2002), uno de los objetivos de la selección de tareas consiste en la generalizabilidad; por lo tanto, entre más "real" sea una tarea, menos generalizable será y mayores riesgos correremos al establecer inferencias de lo que el sustentante es capaz de hacer. Sin embargo, también es importante delimitar el alcance de una tarea en cuanto a las habilidades de escritura que un examinado demuestra para resolver la tarea, puesto que como bien menciona Bachman (2002), predecir futuros desempeños resulta riesgoso.

Entonces, aunque Weigle (2002) menciona que las tareas propuestas para los exámenes de lengua no son reales, esto no significa que sean inauténticas. La diferencia fundamental estriba en que se pretende evaluar los conocimientos lingüísticos de un sustentante y la manera mediante la cual hace uso de ciertas habilidades y estrategias de escritura que son frecuentemente empleadas en el ámbito académico; por ejemplo, como la capacidad de argumentar. Consecuentemente, en contraposición a lo que afirma Weigle (2002), una tarea de examen puede ser auténtica, pero no significa que sea real dentro del contexto para el cual se evalúa.

Más aún, Bachman y Palmer (1996 citados en Bachman 2002: 460) sugieren tres razones por las cuales las tareas del "mundo real" no son apropiadas en ámbitos que son muy amplios: En primer lugar, no todas las tareas relacionarán las habilidades que deseamos evaluar. En segundo lugar, algunas tareas son imprácticas para administrarse en un examen. Y en tercer lugar, no todas las tareas reales serán apropiadas para todos los sustentantes, sobre todo si estas presuponen conocimiento previo del tema o del desarrollo específico de la tarea. De hecho, como apunta Mendoza (en prensa) ¿cómo elegir entre una de las diferentes tipologías comúnmente empleadas a nivel universitario cuando resulta difícil sistematizar los géneros discursivos y las tareas de escritura que les son solicitadas a los estudiantes universitarios?

En aras de lograr mayor autenticidad en el diseño de los ELAA, La inclusión de tareas de escritura (conocidas también como pruebas directas) ha aumentado considerablemente y también ha desplazado en gran medida los mecanismos que evalúan las habilidades de escritura mediante modelos psicométricos (conocidos como pruebas indirectas). Estos últimos, generalmente bajo la modalidad de opción múltiple, fueron ampliamente utilizados durante la segunda mitad del siglo XX debido a su practicidad y facilidad de corrección. Sin embargo, el advenimiento del enfoque comunicativo de los 80 (Canale y Swain 1980, Canale 1983) favoreció cambios importantes en el diseño de exámenes de lengua. Más adelante, en los 90, la necesidad de diseñar exámenes que reflejaran la realidad y las condiciones de uso en contextos específicos (Bachman y Palmer 1996) aceleró el florecimiento de dichas tareas de escritura.

En el caso del IELTS (conocido hasta 1989 como ELTS), por ejemplo, se incluyó en 1980 una tarea de escritura integrada en la cual los candidatos debían remitirse a uno de los textos, previamente leídos en la sección de comprensión de lectura, para componer su redacción. Tras diversas investigaciones conducidas con el fin de validar la prueba, entre las que destacan la de Alderson y Clapham (1993), en 1995 se decidió eliminar la tarea integrada debido al sesgo en la evaluación que se presentaba al mezclar la habilidad de lectura con la escrita, y se optó por una prueba escrita compuesta por dos tareas discretas (una con información visual –generalmente a través de gráficas– y otra con una consigna que requiere de la argumentación u opinión para su elaboración).

La contraparte norteamericana del IELTS, el TOEFL PBT (*Paper-Based Test*), adoptó por vez primera en 1986 una tarea de escritura a modo de ensayo. Sin embargo, dicha prueba era independiente del examen y solamente se administraba en casos en los cuales era solicitada por las instituciones. La prueba escrita del TOEFL se convirtió en obligatoria hasta 1998 con el nacimiento del TOEFL CBT (*Computer-Based Test*). No obstante, este examen no fructificó y rápidamente fue remplazado por la nueva versión denominada TOEFL iBT (*Internet-Based Test*) en 2005. Esta nueva versión incluye una tarea de escritura integrada con comprensión de lectura y comprensión auditiva, y una tarea discreta a manera de ensayo. Sin embargo, una de las mayores dificultades de estas radica en su extensión, puesto que el supuesto trasfondo de que un estudiante es capaz de integrar el material proveniente de dos fuentes (una oral y otra escrita) y condensarlo en un texto de 150 palabras es poco plausible de su capacidad para construir conocimiento a nivel universitario. En este sentido, como bien mencionan Kane et al. (1999), establecer el balance entre estandarización y autenticidad no es fácil. En el caso del contexto hispanohablante, no existe aún ningún ELAA que sea administrado y reconocido a nivel nacional o internacional.

1.1 LOS TIPOS DE TAREAS DE ESCRITURA PARA LOS ELAA

La selección de las tareas y las variables intrínsecas a que estas están sujetas han suscitado una serie de controversias respecto de su autenticidad y pertinencia, pero sobre todo, ha generado cuestionamientos sobre su confiabilidad y validez. Debido a la problemática que representa la concepción, diseño y selección de tareas de escritura para evaluar las habilidades de escritura en un contexto académico, es imprescindible diseñarlas adecuadamente y llevar a cabo pilotajes con el fin de determinar su pertinencia antes de emplearse como mecanismo de evaluación. Bachman (2002) agrega que en la etapa del diseño se debe describir claramente las especificaciones de lo que se pretende evaluar mediante la expresión escrita, por ejemplo, en aras de posteriormente poder seleccionar tareas que sean auténticas. Weigle (2002), por su parte, añade que las consideraciones más importantes que deben tomarse en cuenta para el diseño de tareas escritas son: a) la descripción del dominio de escritura, b) diseñar consignas claras y precisas que sean interpretadas por los sustentantes de la misma forma; es decir, sin ambigüedades ni dificultades léxicas o sintácticas que entorpezcan la comprensión de la tarea y c) uniformidad en los mecanismos de evaluación que permitan obtener resultados consistentes en evaluaciones posteriores o al ser revisados por diferentes correctores.

1.2 LAS VARIABLES INTRÍNSECAS AL TIPO DE TAREA DE ESCRITURA

Las variables que influyen en el desempeño de la tarea son múltiples. Bachman (2002) y Bachman y Palmer (1996: 65) mencionan que al momento de resolver una tarea entran en juego varios factores como los personales (edad, lengua materna, nivel de educación, familiarización con el examen, conocimiento del mundo, conocimiento de la lengua) y los contextuales, generalmente determinados por las características de la tarea. Sin embargo, afirman que hay un proceso de interacción en el cual se ven involucrados esquemas socioafectivos y también el uso de estrategias metacognitivas para su resolución. Sin embargo, dado que el objetivo de esta investigación consiste simplemente en ver qué tan pertinentes son las tareas de escritura seleccionadas respecto del dominio para el cual se desea evaluar, solamente se abordarán las variables relacionadas con la tarea. A este respecto, Weigle (2002) menciona que las variables más representativas que inciden en el diseño de las tareas de escritura son: la posibilidad de elegir entre varias tareas, el contenido, el conocimiento previo del

tema, el uso de material de estímulo, el modo discursivo –mismo que incluye el género textual, la retórica, el patrón de exposición y las demandas cognitivas–, la longitud del texto, el tiempo de respuesta, la consigna, el modo de transcripción y el uso de diccionarios.

1.3 LA SELECCIÓN DE LA TAREA

Incuestionablemente, la decisión más importante consiste en seleccionar la tarea o tareas a través de las cuales se evaluará la habilidad escrita. Bachman (2002) afirma que la dificultad de esta no reside en la tarea *per se*, sino que más bien se encuentra determinada por el proceso de interacción que se lleva a cabo entre el sustentante y la tarea. Entonces, quizás lo que habría que considerar al momento de diseñar tareas de escritura es que estas fomenten la interacción y puedan eficazmente establecer un vínculo socioafectivo en el sustentante y la tarea.

En general, podríamos decir que las tareas de escritura de los ELAA se dividen en dos: discretas (miden una sola habilidad) e integradas (relacionadas alguna otra habilidad). Ejemplos de tareas discretas serían: escribir a partir de un estímulo visual (gráficas o esquemas de procedimiento) o escrito (consignas para redactar ensayos y refranes o frases para buscar su interpretación y postura). En el caso de las tareas integradas, podemos encontrar aquellas que incluyen información escrita y/u oral. Aunque la selección de una tarea de escritura pareciera una decisión sencilla de tomar, en realidad no lo es, ya que la gran crítica que ha recibido toda esta amplia gama de posibilidades reside justamente en que cada una de ellas involucra variables adicionales que interfieren directamente con el nivel de dificultad de la tarea; generando posibles sesgos en los resultados obtenidos.

1.3.1 LAS TAREAS DISCRETAS

Debido a que las tareas discretas involucran la redacción de un tema a partir de una consigna dada, la mayor controversia emerge al seleccionar tareas que involucran o no el conocimiento previo del sustentante. Autores como He y Shi (2012) mencionan que el contenido dado por el tema debe ser tan justo como sea posible, sin que este favorezca a un grupo específico. Hamp-Lyons (1993), Lee y Anderson (2007) comentan que esta desventaja se incrementa aún más cuando se trata de estudiantes de posgrado, dado que muchas áreas de estudio son por naturaleza interdisciplinarias. Más aún, Lee (2008) identificó que no había diferencias sustanciales en los resultados de los estudiantes que escribieron sobre un tema general y otro relacionado con su área general de estudio; por lo cual, el autor sugiere que por practicidad sería mejor elegir un tema general, pues la otra opción podría afectar la manera en que los evaluadores califican. Por ello, Weigle (2002) comenta que debido a las múltiples variables relacionadas con el candidato (antecedente académico, estilo de aprendizaje, edad, sexo, lengua materna, conocimiento sobre el tema, grado académico y tiempo de residencia en el país) lo mejor será excluir el conocimiento del mismo, razón por la cual las pruebas que no evalúan el conocimiento previo encajan bien en los exámenes de alto impacto. Sin embargo, lo anterior ha suscitado controversias en cuanto a la selección de la tarea; puesto que dentro de la escritura académica universitaria se busca que estos no solo lo hagan de manera coherente y siguiendo las normas convencionales de escritura, sino que además sea de manera responsable (citando a las fuentes y evitando el plagio). Entonces, agregan Leki, et al. (2008), una de las críticas más fuertes contra los ELAA es que las tareas que generalmente se plantean eliden completamente el uso de fuentes externas.

Algunos de los problemas que se han evidenciado con la inclusión de tareas generales se deben a la resistencia que en ocasiones demuestran los examinados para resolver la tarea. Petersen (2009) llevó a cabo un estudio en el cual identificó que de un total de 893 estudiantes que respondieron un ensayo de tipo argumentativo, 36 de ellos (4%) mostraron oposición y resistencia abiertamente con frases y títulos que explícitamente retaban el tema para redactar. A este respecto, valdría la pena matizar que quizás la oposición abierta de escribir sobre un tema no reside en la dificultad de la tarea, sino más bien en la falta de elementos en la consigna que estimulen el esquema socioafectivo del estudiante para interactuar apropiadamente con la tarea. Kormos (2012: 394) agrega que “los estudiantes que se encuentran intrínsecamente motivados se involucran en el proceso de aprendizaje porque encuentran el proceso de escritura interesante y placentero, mientras que los estudiantes extrínsecamente motivados llevan a cabo una actividad de aprendizaje con el fin de obtener una recompensa o evitar un castigo”. Por lo tanto, las tareas deberían estimular el factor motivacional del examinando tanto en contextos académicos como en exámenes de dominio para el ámbito académico. A este respecto, Lee (2008) menciona que los estudiantes prefieren elegir entre varias temáticas, pues además de incrementar su motivación, los estudiantes buscan temas con los cuales tengan mayor familiaridad y destreza para escribir.

La última variante de tareas discretas es aquella en la cual el estímulo que se proporciona para escribir es visual. Este tipo de información generalmente se da a través de imágenes que denotan cómo se lleva a cabo un procedimiento o gráficas para describir y contrastar. En estos casos, el estímulo visual presenta el contenido *per se*, lo cual permite que el candidato

recurra a su conocimiento lingüístico y léxico para describirlo. Weigle (2002) arguye que la carencia de estímulo escrito u oral permite que el candidato explote al máximo su conocimiento del mundo y sus recursos lingüísticos. Sin embargo, al seleccionar este tipo de tareas se debe tener la suficiente cautela de escoger elementos visuales que sean lo suficientemente comprensibles y claros, así como con contenido que no sea especializado, ya que esto pondría en desventaja a quienes carecieran del conocimiento léxico sobre el tema en cuestión. Sin embargo, como bien apunta Yang (2012), las tareas integradas con material visual estimulan el uso de determinado léxico que representa un reto mayor que la dificultad sintáctica. Uysal (2010), en la revisión crítica que hace del IELTS académico, comenta que las tareas más simples con menos información visual permiten un mejor desempeño de los candidatos.

El problema de las gráficas es que estas presentan diversas variables intrínsecas: el formato de la gráfica (barras, pastel, lineal) y la familiaridad que puedan tener los sustentantes con estas (Shah y Freedman 2009). Estos autores concluyen la familiarización con los datos permite establecer mejores inferencias, además de que quienes poseen altas habilidades con la descripción de gráficas pueden interpretarlas mejor. Por eso, Yang (2012) afirma que el uso de estrategias es primordial para la consecución de la tarea, por lo cual debería hacerse énfasis en los cursos de preparación la inclusión de este tipo de tareas integradas, así como de las estrategias requeridas para abordarlas. Aunado a lo anterior, como afirman Shah y Freedman (2009) y Hyland (2009), la importancia de trabajar con gráficas en el ámbito académico es esencial.

1.3.2 LAS TAREAS INTEGRADAS

Las tareas integradas consisten en escribir un texto a partir de información escrita y/u oral que se da como input. Este tipo de tareas han sido incluidas recientemente en los exámenes de lenguas con fines académicos con miras a dar cuenta sobre la manera mediante la cual los candidatos construyen el conocimiento a partir de la inclusión de información proveniente de diversas fuentes; por lo tanto, resultan “más auténticas” para el contexto para el cual se desea evaluar: el académico. El supuesto trasfondo reside en que las tareas que los estudiantes llevan a cabo en la universidad incorporan la comprensión de lectura y oral; mientras que las tareas independientes se encuentran aisladas de estas habilidades. Plakans (2010) menciona que una cuestión fundamental para la resolución de una tarea escrita reside en la representación que el alumno se hace dependiendo del tipo de tarea que requiere resolver. Entonces, cuando un estudiante identifica la clase de producto que se espera de la escritura, este atraviesa por un proceso de entendimiento en el cual pueda identificar el tipo de habilidades y procesos que requerirá para llegar a este producto determinado: el tipo de tarea escrita solicitada.

Cumming (2013: 2), en el análisis que efectúa de las tareas integradas, menciona que las ventajas de estas son: tareas más auténticas, escritura con contenido “responsable” y evaluación de las habilidades lingüísticas en la construcción-integración del conocimiento. Sin embargo, las desventajas que evidencian son: mezcla de otras habilidades en la medición de la escritura, involucramiento de géneros que no están bien definidos y que son difíciles de evaluar, requerimiento de ciertas habilidades mínimas para llevarla a cabo, y la inclusión de información que es difícil determinar si su procedencia subyace en la capacidad de producción del sustentante.

Dentro de las dificultades que encarna la implementación de tareas integradas se halla precisamente la incorporación de habilidades orales y de lectura que, en caso de no estar bien consolidadas, dificultarán la producción escrita de los estudiantes (Leki et al. 2009). De hecho, en el estudio que Plakans (2010) condujo para ver de qué forma la representación de la tarea (discreta o integrada) impactaba en el proceso de escritura de los examinados, la autora comenta que la incorporación de textos en las tareas integradas puede servir de apoyo para algunos estudiante, aunque otros pueden verse desfavorecidos por la complejidad del tópico o por la dificultad que entraña la construcción del conocimiento al incorporar el material de lectura en el escrito solicitado. Aunado a lo anterior, una dificultad más se encuentra en determinar qué fue copiado literalmente del texto, tarea que les resulta sumamente compleja de llevar a cabo a los evaluadores humanos (Cumming 2013; Cumming, Kantor, Baba, Erdosy, Eouanzoui y James 2005).

A este respecto, Weigle y Parker (2012: 128) identificaron que el porcentaje tomado literalmente del texto era en la mayoría de los casos menor al 8%, además de que este se refería normalmente a colocaciones. Sin embargo, se evidenció que los estudiantes con puntajes bajos tienden a citar porciones más largas provenientes del texto de entrada, a diferencia de los estudiantes con mayor dominio de la lengua, quienes parafrasean más. Haciendo eco al estudio de Weigle y Parker (2012), el de Plakans y Gebril (2013) confirmó que el uso apropiado de las fuentes incide positivamente en la escritura. Los autores identificaron que la tarea integrada favorece la inclusión del contenido del audio, mientras que el del texto se emplea solamente para contrastar la información escuchada.

No obstante, las tareas integradas del TOEFL iBT que incluyen un texto de entrada escrito y otro auditivo nos llevan al cuestionamiento de si en realidad la comprensión auditiva, por lo menos dentro del contexto universitario, conlleva tanto peso al momento de escribir. Por lo tanto, sería importante indagar si las prácticas de escritura universitaria buscan la integración de estas dos habilidades; de lo contrario, valdría la pena considerar si una tarea integrada con dos o múltiples fuentes

escritas sería más apropiada que una que incluye comprensión de lectura y auditiva. Siguiendo la línea de Weigle y Parker (2012), ¿cómo distinguir del material tomado del texto cuando este no ha sido citado? Aunado a lo anterior, ¿se puede evidenciar en un par de textos y un corpus de 150 palabras la habilidad de integrar el material proveniente de diversas fuentes? Sin lugar a dudas, como asevera Hyland (2009), el correcto uso del material escrito con su correcta acreditación al autor es un componente importante de la escritura académica –misma que debería figurar dentro del constructo de los exámenes–; sin embargo, la gran interrogante que subyace es si en verdad esta habilidad puede evaluarse en un texto de 150 palabras con una o dos fuentes de entrada.

Finalmente, dentro del contexto hispanohablante, así como en el de la mayoría de las demás lenguas (excepto, por supuesto, del inglés), una gran cantidad del material que se emplea para construir el conocimiento escrito procede de fuentes como el inglés. Por lo tanto, valdría la pena indagar tres cosas: en primer lugar, ¿qué porcentaje de las lecturas solicitadas por los docentes universitarios provienen de lenguas extranjeras o indígenas y cuáles? En segundo lugar, ¿qué porcentaje de las lecturas que hacen los alumnos es en otra lengua diferente del español? Y por último, ¿hasta qué punto sería óptimo emplear tareas integradas con textos escritos en español en disciplinas en donde casi el 100% del material de lectura está en inglés? No obstante, pese a las dificultades y retos que representa el uso de tareas integradas, la incorporación de estas en exámenes de lengua con fines académicos merece atención especial.

1.4 OTRAS VARIABLES IMPORTANTES

Existen otras variables que inciden indirectamente en la prueba como: el modo discursivo, la claridad en la consigna, el tiempo de respuesta, el modo de transcripción y el uso de diccionario. No obstante, pese a que los hallazgos respecto de estas variables no han logrado dilucidar cuál sería la mejor manera abordar dichas variables, resulta importante considerar su relevancia.

1.4.1 EL MODO DISCURSIVO

Algunos detractores de la autenticidad de las pruebas escritas de los exámenes con fines académicos (Moore y Morton 2005, Uysal 2010) arguyen que estas son diferentes a las tareas que se llevan a cabo en un contexto real. Sin embargo, como bien afirma Bachman (2002), la autenticidad no significa que una tarea sea real, sino que esta se encuentre relacionada con el uso de la lengua requerido en el contexto para el cual se desea evaluar; en este caso, el académico. Por supuesto que es natural que los ensayos que se redactan en un examen con fines académicos sean diferentes a los ensayos que los estudiantes llevan a cabo en su vida universitaria. Por consiguiente, estudios como los que tratan de encontrar la correlación (Chan, Wu y Weir 2014) o falta de esta (Moore y Morton 2005) entre el texto redactado en las tareas de un examen de certificación con fines académicos y textos creados dentro del contexto universitario son irrelevantes, puesto que evidentemente se trata de tareas que intrínsecamente son diferentes.

1.4.2 LA CARIDAD EN LA CONSIGNA

Por ejemplo, en el caso de la consigna, como bien sugiere Weigle (2002) esta debe ser clara y concisa, puesto que las consignas muy elaboradas que en ocasiones incluyen ejemplos son sujetas a ser copiadas por los estudiantes. Por el contrario, cuando la consigna no es muy clara o demasiado abierta a interpretación, no permite delimitar la tarea y los estudiantes tienden a escribir sobre una vasta variedad de tópicos. En estudios que se han conducido para identificar las secuencias formulaicas o paquetes léxicos se ha evidenciado que los estudiantes menos cualificados usan una mayor cantidad de estas secuencias (Staples, Egbert, Biber y McClair 2013) y también copian más paquetes léxicos provenientes de las consignas o, en el caso de las tareas integradas, de los textos. Kobrin, Deng y Shaw (2011) mencionan que la redacción de la consigna determina la dificultad de la tarea. Dicha variable se percibe, por ejemplo, cuando hay palabras o calificadores de una afirmación, cuando se expresa un punto de vista neutral o con una opinión, así como con argumentos con una o dos posturas. Incluso, una consigna demasiado larga podría afectar el entendimiento de aquellos estudiantes con menos dominio lingüístico.

1.4.3 EL USO DE DICCIONARIO

Finalmente, otra de las discusiones se ha centrado en los beneficios del uso de un diccionario bilingüe o monolingüe para la resolución de las tareas escritas. Weigle (2002) comenta que el uso de diccionarios en las pruebas de lengua cronometradas no se permite debido a que el vocabulario es parte de lo que se pretende evaluar. La autora agrega que el uso de diccionarios tampoco es una habilidad que todos los sustentantes puedan desplegar satisfactoriamente, ya que dicha

habilidad se restringe mayoritariamente a quienes son considerados como “buenos escritores”. En general, en la mayoría de los ELAA no se permite el uso de estos; sin embargo, la versión alemana del examen académico, TEST DAF, sí permite su uso. El estudio que East (2006) condujo con respecto al uso del diccionario partió del supuesto que el uso de un diccionario correspondería a una muestra más auténtica en los exámenes cronometrados, puesto que una de las críticas más asiduas a este tipo de exámenes se debe justamente a que la falta de recursos genera una situación “artificial” de escritura que no se equipara con lo que un estudiante redacta en un contexto universitario.

De los hallazgos más sobresalientes del estudio que East (2006) llevó a cabo, se destaca que el uso del diccionario incrementó considerablemente la sofisticación del léxico³ en los participantes. Sin embargo, una segunda mirada hacia la precisión en el uso del vocabulario demostró aproximadamente un 50% de errores en la selección de la palabra adecuada al contexto en el caso de los sustentantes con menores habilidades de escritura. Por lo tanto, si bien es cierto que el léxico fue más sofisticado, también es cierto que la mitad de este fue seleccionado inapropiadamente, por lo cual su uso no proporcionaba ningún valor agregado a la escritura. En contraparte, los estudiantes con mayores habilidades de escritura si se vieron beneficiados por el uso del diccionario, ya que las palabras que seleccionaron fueron apropiadas para el contexto. El autor concluye, siguiendo la línea de Weigle (2002), que el uso de diccionarios en los ELAA conlleva un impacto negativo que incluso sobrepasa el beneficio de que los examinandos escriban con un léxico más sofisticado, pero no siempre pertinente al contexto.

1.4.4 EL MODO DE TRANSCRIPCIÓN

Una de las grandes interrogantes que se presenta hoy en día al evaluar los exámenes cronometrados reside justamente en por qué se hace a mano cuando los estudiantes escriben a computadora. Weigle (2002) comenta que los correctores tienden a evaluar con mayor dureza los textos transcritos a computadora por varias razones: la primera, porque los textos se ven más cortos que los escritos a mano, segundo, porque no es posible evidenciar la competencia estratégica de la edición del texto (como podría ser cambiar una oración de lugar o borrar una palabra para buscar una más apropiada), y en tercer lugar, porque estamos acostumbrados a visualizar los textos escritos a computadora como productos finales, mientras que los redactados a mano son frecuentemente concebidos como borradores. Horkay, Bennet, Allen, Kaplan y Yan (2006) también identificaron mejores resultados, aunque no estadísticamente significativos, en los textos transcritos a mano, aun cuando los autores aducen esta diferencia a la familiaridad que los estudiantes presentan para escribir a computadora.

Mogey y Hartley (2013) reconocen estas mismas disimilitudes al momento de evaluar trabajos escritos a mano o a computadora, pero resaltan, por un lado, la importancia de generar conciencia en los estudiantes y correctores de qué tan largo es un ensayo escrito a computadora en comparación con aquellos escritos a mano. Por otro lado, destacan la necesidad de que los correctores se familiaricen con los textos escritos a computadora, vistos no como un trabajo final, sino más bien como un borrador que puede presentar errores. Más aún, los autores aseveran que en un futuro cercano, si no ya existente, los estudiantes redactarán los ensayos a computadora, incluso en los exámenes, y que los ensayos escritos a mano desaparecerán (Mogey y Hartley 2013: 91).

Contradictoriamente al paulatino proceso de desaparición de la escritura a mano, en el estudio que Mogey, Cowan, Paterson y Purcell (2010) condujeron para ver qué modo de transcripción escogería un estudiante si se le diera la oportunidad de hacerlo, los autores identificaron que una gran mayoría optó por la redacción a mano (solamente 16 de 204 candidatos prefirieron hacerlo a computadora). Al parecer algunas de las posibles causas en elegir la escritura a mano se deben a la posible dificultad de escribir a computadora, a la desconfianza con la tecnología y también a que se trataba de estudiantes que recién habían ingresado a la universidad, quienes pudieron haber visualizado la escritura a computadora como “demasiado riesgosa”. Pese a la contundente preferencia en el modo de transcripción, los autores concluyen que la posibilidad de elegir el modo de transcripción debería quedar al arbitrio de los examinandos.

1.5 HIPÓTESIS

Siguiendo la línea de Kane (2002b, 2006, 2013) para validar exámenes y siguiendo el modelo argumentativo de Toulmin (1958), se elaboraron las siguientes hipótesis para llevar a cabo el proceso de validación de la selección de las pruebas escritas del EXELEAA.

Las tareas concebidas para la prueba escrita del EXELEAA son adecuadas para evidenciar las habilidades de escritura requeridas a nivel universitario.

H2

Los temas planteados en las tareas de la prueba escrita diseñadas para el EXELEAA facilitan la interacción entre esta y los examinados.

2. MÉTODO

El estudio se dividió en dos fases: pilotajes y primera aplicación del examen. En el pilotaje se emplearon dos muestras diferentes: a) alumnos de posgrado ya inscritos en la universidad y b) alumnos de licenciatura recién llegados a México para cursar un semestre o año académico en la universidad. Para la primera aplicación del examen, se trabajó con estudiantes de posgrado inscritos en la universidad.

2.1 PARTICIPANTES

2.1.1 PRIMERA FASE: PILOTAJES

2.1.1.1 PRIMER PILOTAJE

A finales de noviembre de 2013 se convocó a toda la población de posgrado (maestría y doctorado) no hispanohablante de la Universidad Nacional Autónoma de México que se encontraba realizando sus estudios en el período de otoño de 2014. Debido a que el total era de 119 alumnos, contando alumnos de primer ingreso y reingreso, se decidió convocar a toda la población. De esta, 21 sujetos participaron en este primer pilotaje. Los estudiantes pertenecían a 9 distintas nacionalidades y 14 distintos programas de posgrado. A quienes participaron en el pilotaje se les otorgó una constancia y también se les ofreció un taller gratuito de escritura académica.

2.1.1.2 SEGUNDO PILOTAJE

Debido a que la primera muestra que se obtuvo no fue tan amplia como se esperaba. En enero de 2014 se convocó a una segunda sesión con los estudiantes de nuevo ingreso al posgrado. Sin embargo, debido a que la población estudiantil de posgrado que ingresa en los semestres de invierno es tan escasa, ningún estudiante de posgrado se inscribió a dicha sesión. Sin embargo, debido al interés de administrar dicha prueba a los estudiantes de licenciatura no hispanohablantes que vienen a México a cursar un semestre o año académico como parte de sus estudios universitarios, se convocó a una sesión para el pilotaje con esta población. Con la colaboración de la Dirección General de Cooperación e Internacionalización (DGEI), participaron 18 estudiantes recién llegados a México a cursar un semestre o año de intercambio académico en la UNAM en el período de otoño de 2014. Los alumnos pertenecían a 7 distintas nacionalidades y a 11 distintas disciplinas académicas. Nuevamente, quienes participaron en el pilotaje se les otorgó una constancia y también se les ofreció un taller gratuito de escritura académica.

2.1.2 SEGUNDA FASE: APLICACIÓN DEL EXAMEN

Finalmente, una vez que se llevaron a cabo los pilotajes y se llevaron a cabo enmiendas a las tareas escritas propuestas, se invitó nuevamente a toda la población no hispanohablante de posgrado inscrita en el semestre de otoño de 2014 a que participara en el examen, mismo que se ofreció de forma gratuita y con una constancia oficial, pues muchos de los estudiantes ya cuentan con el requisito de lengua antes de iniciar su posgrado. El examen se administró a principios de diciembre de 2014, y de los 132 estudiantes convocados, participaron 21 alumnos de 11 distintas nacionalidades y 12 disciplinas académicas.

2.2 MATERIAL

De acuerdo con el diseño del constructo para la prueba escrita, se eligieron dos tareas: una descripción y contraste de gráficos, y una redacción de un ensayo argumentativo. Para los pilotajes se emplearon las pruebas escritas que figuran en el cuadro 1, mientras que para la primera aplicación del examen, se emplearon las pruebas escritas que aparecen en el cuadro 2.

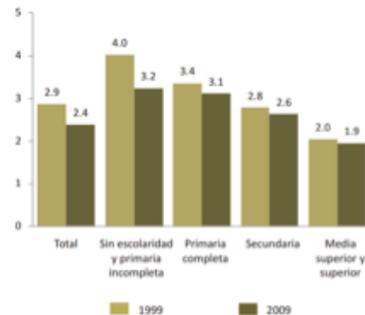
<p>Tarea 1</p> <p>A continuación se presentan dos gráficas sobre el logro educativo que han alcanzado en la última década las hijas con respecto a sus padres en un país de América Latina. En el cuadro de texto que aparece debajo de las gráficas, descríbelas y contrástalas utilizando mínimo 150 palabras. Toma 25 minutos para esta tarea.</p>	
 <p>NIVEL EDUCATIVO DE LOS PADRES EN 2008</p>	 <p>NIVEL EDUCATIVO DE LAS HIJAS EN 2008</p>
<p>Tarea 2</p> <p>a) Lee la siguiente información: "Si bien la educación por sí misma amplía los horizontes de los individuos y constituye un beneficio para éstos, la situación educativa de hombres y mujeres tiene diferentes beneficios para unos y otros y para sus hijos e hijas" (Milosavljevic, 2007, p.96).</p> <p>b) En el cuadro de texto que aparece en la parte inferior, escribe un ensayo sobre el tema de la educación de las mujeres. Puedes retomar información de la Tarea 1.</p> <ul style="list-style-type: none"> - Menciona la importancia del tema. - Establece una postura al respecto. - Establece los pros y los contras de la temática. - Menciona lo que ocurre en tu país. - Concluye. <p>Escribe mínimo 250 palabras. Tienes 45 minutos para realizar esta tarea.</p>	

Cuadro 1: tareas empleadas para los pilotajes

Tarea 1

A continuación se te presenta una gráfica de barras. En el cuadro de texto que aparece debajo de las gráficas, descríbelas y contrástalas utilizando aproximadamente 200 palabras. Toma 25 minutos para esta tarea.

Promedio de hijos de mujeres mexicanas



Fuente: INEGI. XII Censo General de Población y Vivienda 2000. Base de datos; Censo de Población y Vivienda 2010. Base de datos de la muestra.

Tarea 2

En algunos países existen solamente sistemas de salud pública, mientras que en otros existen sistemas públicos y privados. Con base en dicha temática, escribe un ensayo argumentativo con base en lo siguiente:

- Menciona la importancia del tema.
- Establece una postura al respecto.
- Establece los pros y los contras de los sistemas públicos y los privados.
- Menciona cómo es el sistema de salud pública en tu país.
- Concluye.

Escribe mínimo 300 palabras. Tienes 45 minutos para realizar esta tarea.

Cuadro 2: tareas empleadas para la primera aplicación

2.3 PROCEDIMIENTO

En primer lugar, con base en la literatura respecto del alcance de las pruebas discretas e integradas, así como de un estudio previo sobre las habilidades de escritura requeridas a nivel universitario (Mendoza en prensa), se estableció el alcance de las tareas aquí propuestas. Posteriormente, con el fin de identificar posibles dificultades que hubieran encontrado los estudiantes en relación con las gráficas o el tema argumentativo, se hizo un análisis del corpus que pudiera dar cuenta de estos problemas. Asimismo, para evidenciar la copia literal de paquetes léxicos, se identificaron las pequeñas muestras dadas en las consignas para ver en qué medida los estudiantes copiaban información literal proveniente de estas. En el examen no se permite emplear diccionarios, pues se pretende que el sustentante explote al máximo sus recursos léxicos. Finalmente, el modo de transcripción de los pilotajes fue a mano, mientras que para la primera aplicación se hizo a máquina. Esto último con el fin de poder evidenciar diferencias en la percepción del evaluador al corregir uno u otro modo.

3. RESULTADOS

A continuación se reportan los rasgos más sobresalientes de las dos tareas discretas incluidas en la primera fase del EXELEAA: descripción y contraste de gráficas, y redacción de un ensayo argumentativo. En el caso la gráfica, el primer problema que identificamos se debía a la interpretación de la gráfica, puesto que varios estudiantes escribieron que los datos no se podían comparar debido a que no se contaba con la información de los hijos, lo cual impedía que se pudiera contrastar eficazmente el logro educativo de las mujeres con sus padres.

Ejemplos de problemas evidenciados para contrastar las gráficas.

- “De manera general, se observa un incremento del nivel de educación entre estas dos generaciones. Pero no nos podemos dar cuenta de la verdadera diferencia entre los hombres y las mujeres, dado que no sabemos la proporción de hijos escolarizados, etc.”⁴.
- “Para poder hacer comparaciones de género, se requieren estadísticas tocante a los hijos de estos padres”.

En segundo lugar, cabe mencionar que algunos estudiantes retomaron paquetes léxicos de la consigna, tanto de las gráficas como del ensayo; aun cuando este input era mínimo. Esto demuestra que cuando el estudiante posee fuentes escritas es frecuente que copie fragmentos del texto, lo cual resulta particularmente difícil de identificar para un evaluador humano.

Ejemplo de calcos literales de las tareas 1 y 2.

- “Vimos dos gráficas sobre el logro educativo que han alcanzado en la última década las hijas con respecto a sus padres en un país de América Latina”⁵.
- “En estos dos gráficos recibimos información sobre el logro educativo que han alcanzado las hijas con respecto a sus padres”.
- “Las gráficas muestran el logro educativo de las hijas con respecto a sus padres en un país de América Latina”.
- [Milosavljevic dijo “Si bien la educación por sí misma amplía los horizontes de los individuos y constituye un beneficio para éstos, la situación educativa de hombres y mujeres tiene diferentes beneficios para unos y otros y para sus hijos e hijas”].
- [Según Milosavljevic (2007: 96) “Si bien la educación por sí misma amplía los horizontes de los individuos y constituye un beneficio para éstos”].

Otra de las dificultades observadas en el ensayo argumentativo propuesto se debió a que en realidad la consigna no permitía que los estudiantes utilizaran argumentos en contra de la educación de las mujeres, dado que esto representaría una postura desigual de género. Sin embargo, dentro de la consigna se especifica que mencionen argumentos en contra y en la rúbrica se evalúa el uso de dichos contraargumentos.

Ejemplo de la problemática de evidenciar contraargumentos:

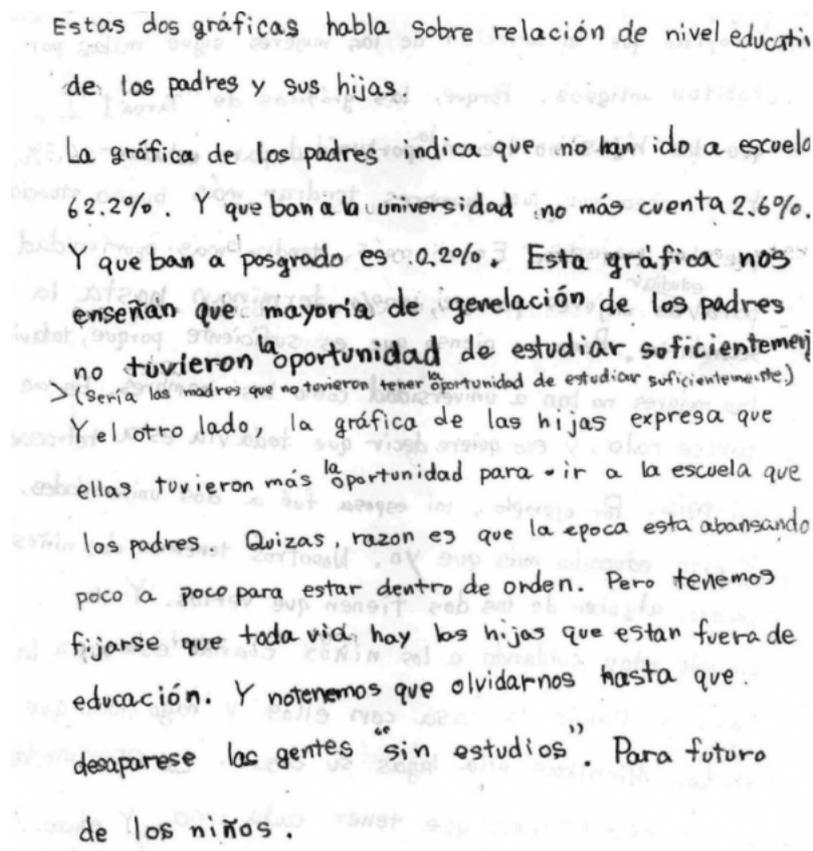
- “Milosavljevic dice aquí que la educación conlleva beneficios diferentes para hombres y mujeres pero yo no veo diferencias de género en los beneficios de la educación”.
- “Tal vez, se dice que la situación educativa tiene diferentes beneficios para los hombres y las mujeres, no obstante, yo considero que la educación es importante para todos”.
- “[...] estoy muy sorprendida cuando se dice que “la situación educativa de hombres y mujeres tiene diferentes beneficios para unos y otros y para sus hijos e hijas” todo el mundo debería tener los mismos resultados y beneficios”.
- “No creo que haya el mínimo argumento en contra del incremento al acceso a la educación para las mujeres.”

La otra problemática que se evidenció en la propuesta inicial de las tareas 1 y 2 es que estas se encontraban ligadas; es decir, ambas tareas versan sobre la educación de las mujeres. Además, en la consigna de la tarea 2 se explicita que el sustentante puede retomar información de la tarea 1. En realidad, ambas tareas deberían ser independientes una de la otra, pues además eso coadyuvaría a que los estudiantes exploten al máximo su repertorio léxico.

Finalmente, cabe destacar que el modo de transcripción a mano dificulta considerablemente la redacción de los textos y en múltiples ocasiones el evaluador requiere interpretar lo que el estudiante ha escrito, aunque también es recurrente que se brinquen los segmentos en los cuales no se puede comprender el texto, lo cual deja huecos las ideas plasmadas en el escrito. En contraparte, es cierto que cuando el texto se transcribe a mano se ve mucho más largo de lo que en realidad es, lo cual visualmente impacta en el evaluador como si el desarrollo hubiera sido más sustancial, aunque en realidad se están

valorando aspectos superficiales relacionados con una mera apreciación visual. Aunado a lo anterior, es cierto que el modo de transcripción se percibe como si fuera un primer borrador, mientras que los errores en computadora destacan más, pues uno está acostumbrado a verlo como producto final. Otro impacto visual es que la letra que es más legible tiende a incidir positivamente en cómo percibe el evaluador el texto de entrada. Pareciera entonces que el contenido del texto es también, por ende, más claro. A continuación se presentan diversos modos de transcripción a mano para ver cuáles son algunos de los problemas que se evidencian con esta manera de escribir.

La caligrafía, puede incidir positiva o negativamente en el evaluador (ver imagen 1). En este caso, quizás la falta de familiarización con el alfabeto grecolatino haría suponer a primera vista que el texto fue escrito por un niño, lo cual puede incidir negativamente en la manera en la cual el evaluador percibe el nivel del sustentante tan solo por su caligrafía. Si bien es cierto que efectivamente el texto era muy básico, el evaluador ya tiene una primera impresión que podría no ser así.



Estas dos gráficas habla sobre relación de nivel educativo de los padres y sus hijas.

La gráfica de los padres indica que no han ido a escuela 62.2%. Y que van a la universidad no más cuenta 2.6%. Y que van a posgrado es 0.2%. Esta gráfica nos enseñan que mayoría de generación de los padres no tuvieron la oportunidad de estudiar suficientemente.

> (Sería los padres que no tuvieron tener la oportunidad de estudiar suficientemente.)

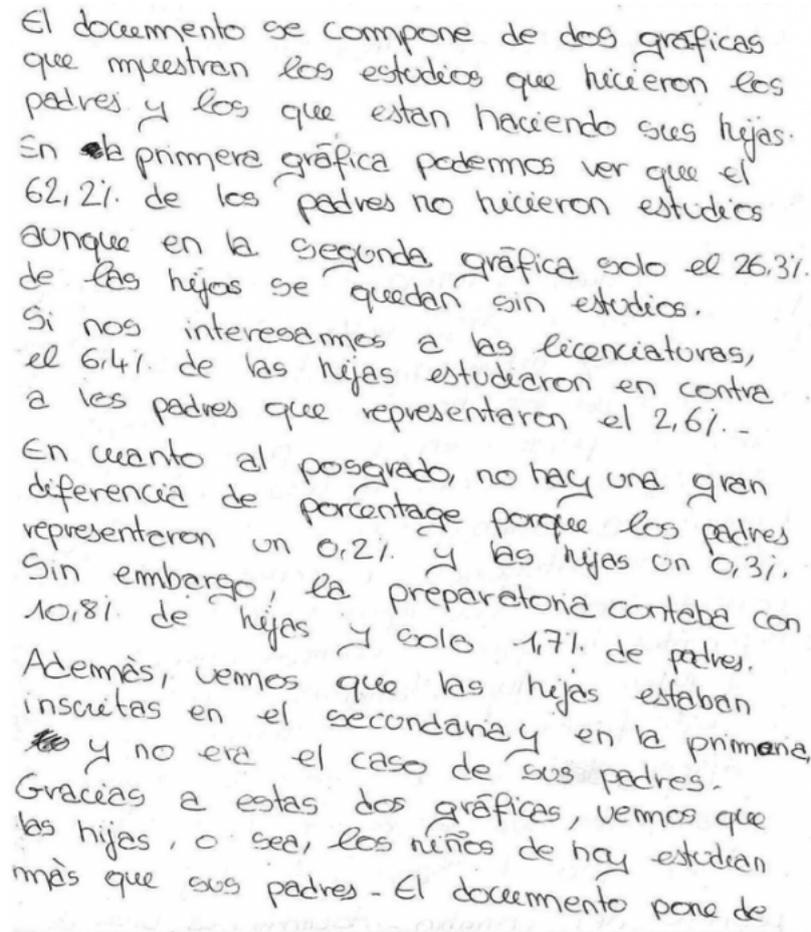
Y el otro lado, la gráfica de las hijas expresa que ellas tuvieron más la oportunidad para ir a la escuela que los padres. Quizás, razón es que la época esta abansando poco a poco para estar dentro de orden. Pero tenemos fijarse que toda vida hay los hijas que estan fuera de educación. Y notenemos que olvidarnos hasta que desaparese las gentes "sin estudios". Para futuro de los niños.

Imagen 1: La mala caligrafía incide negativamente en la percepción del nivel de escritura del sustentante.

Transcripción del texto que figura en la imagen 1

"Estas dos gráficas habla sobre relación de nivel educativo de los padres y sus hijas. La gráfica de los padres indica que no han ido a escuela 62.2%. Y que van a la universidad no más cuenta 2.6%. Y que van a posgrado es 0.2%. Esta gráfica nos enseñan que mayoría de generación de los padres no tuvieron la oportunidad de estudiar suficientemente. Y el otro lado, la gráfica de las hijas expresa que ellas tuvieron más la oportunidad para ir a la escuela que los padres. Quizás, razón es que la época esta abansando poco a poco para estar dentro de orden. Pero tenemos fijarse que todavía hay las hijas que estan fuera de educación. Y notenemos que olvidarnos hasta que desaparese las gentes "sin estudios". Para futuro de los niños" (131 palabras).

El tamaño de la letra influye en la percepción visual del evaluador, puesto que parece mucho más extenso de lo que en realidad es. El texto que aparece en la imagen 2 fue escrito en toda una cuartilla. En el caso opuesto, cuando se revisaron los textos transcritos a computadora, estos se percibían como demasiados pequeños, dado el contraste que se notaba entre los textos transcritos a mano versus a computadora.



El documento se compone de dos gráficas que muestran los estudios que hicieron los padres y los que están haciendo sus hijas. En la primera gráfica podemos ver que el 62,2% de los padres no hicieron estudios aunque en la segunda gráfica solo el 26,3% de las hijas se quedan sin estudios. Si nos interesamos a las licenciaturas, el 6,4% de las hijas estudiaron en contra a los padres que representaron el 2,6%. En cuanto al posgrado, no hay una gran diferencia de porcentaje porque los padres representaron un 0,2% y las hijas un 0,3%. Sin embargo, la preparatoria contaba con 10,8% de hijas y solo 1,7% de padres. Además, vemos que las hijas estaban inscritas en el secundaria y en la primaria y no era el caso de sus padres. Gracias a estas dos gráficas, vemos que las hijas, o sea, los niños de hoy estudian más que sus padres. El documento pone de

Imagen 2: El tamaño de la letra distorsiona la percepción de su longitud.

Transcripción del texto que figura en la imagen 2

“El documento se compone de dos gráficas que muestran los estudios que hicieron los padres y los que están haciendo sus hijas. En la primera gráfica podemos ver que el 62,2% de los padres no hicieron estudios aunque en la segunda gráfica solo el 26,3% de las hijas se quedan sin estudios. Si nos interesamos a las licenciaturas, el 6,4% de las hijas estudiaron en contra a los padres que representaron el 2,6%. En cuanto al posgrado, no hay una gran diferencia de porcentaje porque los padres representaron un 0,2% y las hijas un 0,3%. Sin embargo, la preparatoria contaba con 10,8% de hijas y solo 1,7% de padres. Además, vemos que las hijas estaban inscritas en el secundaria y en la primaria y no era el caso de sus padres. Gracias a estas dos gráficas, vemos que las hijas, o sea, los niños de hoy estudian más que sus padres. El documento pone de...” (181 palabras).

Finalmente, había textos que resultaban difíciles de leer, con lo cual la lectura se veía interrumpida y en ocasiones el evaluador debía interpretar mediante el contexto las palabras o paquetes léxicos que no comprendía (ver imagen 3).

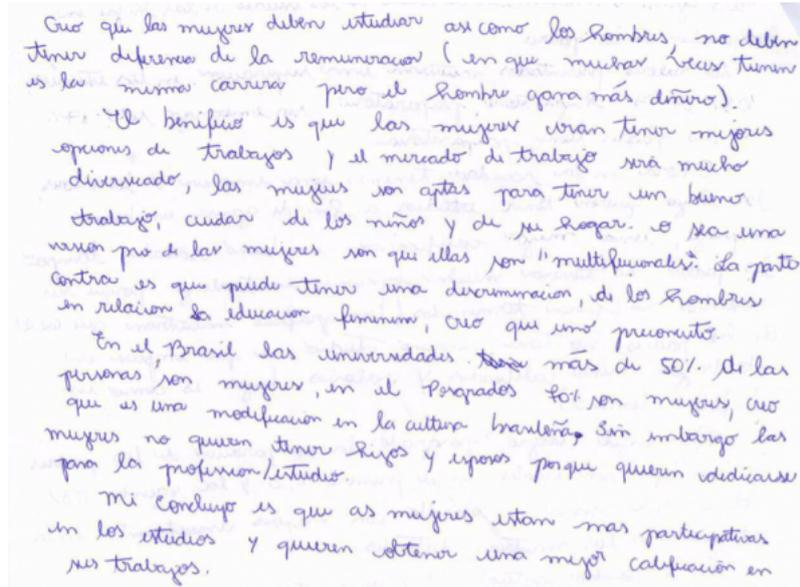


Imagen 3: La caligrafía en ocasiones dificulta la lectura de los textos.

Debido a las problemáticas evidenciadas en las tareas propuesta para el pilotaje, se decidió llevar a cabo modificaciones para la primera aplicación oficial del examen. En primer lugar, se cambió el modo de transcripción y los estudiantes escribieron a computadora. Esto incrementó considerablemente el número de palabras que utilizaron, sobre todo para los ensayos, en los cuales hubo sustentantes que incluso escribieron más de 700 palabras en un texto en el cual se solicitó un mínimo de 300. Esto da cuenta importante de la velocidad que implica no solo la escritura sino también la edición de textos que son escritos a computadora.

En cuanto a la nueva propuesta de la gráfica, se evitó incluir cualquier tipo de información que el sustentante pudiera copiar y calcar en su texto. Sin embargo, nuevamente se presentaron problemas al momento de interpretarlas, dado que cuatro de los 22 estudiantes malinterpretaron la gráfica pensando que esta mostraba que entre mayor era el número de hijos de mujeres mexicanas, estos contaban con menor escolaridad, cuando en realidad la gráfica muestra que las mujeres con mayor logro educativo tienden a concebir menor número de hijos.

En el caso de la segunda tarea, además de que esta se diseñó de forma completamente independiente a la primera, se elidió cualquier información en la consigna que pudiera ser copiada textualmente por parte del sustentante. Asimismo, se buscó un tema que pudiera ser de interés general para los estudiantes y en el cual se pudieran establecer argumentos a favor o en contra de una postura. En este caso, cabe mencionar que la mayoría de los estudiantes mostraron gran interés en la tarea propuesta, lo cual se evidencia en los siguientes ejemplos:

Ejemplos que demuestran la relevancia del tema y, por lo tanto, de la interacción del estudiante con la tarea.

- "El tema de salud pública siempre ha sido **significativa** por abrir la puerta de protección, seguridad y comodidad de cada habitante".
- "En este ensayo intentaré discutir los sistemas de salud públicos y privados en general, en México y en mi país, Alemania. Por las diferencias existentes en los sistemas es un tema corriente y políticamente **mucho importante**".
- "El tema de sistemas de salud es **sumamente importante**, dado que afecta directamente a todos seres humanos".
- "La cuestión de salud siempre es de **suma importancia** para los gobiernos de los países en todo el mundo".
- "El tema de la salud pública es un tema **bastante significativo** hoy en día".

En general, la única problemática que se evidenció fue que dos estudiantes reaccionaron emocionalmente ante el tema y el tema viró de un tono argumentativo a uno narrativo respecto de las malas experiencias médicas que habían vivido con el sistema de salud público o privado en México, cuando en realidad no se les pedía que narraran experiencias personales con los sistemas de salud.

4. DISCUSIÓN

Con respecto a la primera hipótesis, cabe mencionar que al seleccionarse tareas discretas, la posibilidad de evidenciar la construcción del conocimiento queda fuera. Entonces, resulta importante delimitar en primera instancia el alcance de la prueba escrita, ya que como se evidenció en un estudio anterior a este (Mendoza en prensa), los docentes de posgrado de la Universidad Nacional Autónoma de México consideran que las habilidades de escritura primordiales para que los estudiantes universitarios sean académicamente competentes giran fundamentalmente en torno a dos grupos: las habilidades de escritura necesarias para construir conocimiento y las habilidades de escritura de índole académico para comunicar efectivamente dicho conocimiento. Sin embargo, como también se pudo observar en las discusiones manifiestas en este artículo, resulta importante reflexionar en torno a lo siguiente: ¿es posible evidenciar la construcción del conocimiento en un texto de aproximadamente 150 palabras con el input proporcionado por uno o dos textos? Y, en el contexto hispanohablante, ¿siempre construyen los estudiantes el conocimiento de fuentes provenientes del español? Aunado a lo anterior, como se pudo apreciar en este trabajo, incluso cuando las consignas presentan un mínimo de estímulo escrito, este en ocasiones es copiado literalmente por los sustentantes, con lo cual resulta difícil que el evaluador humano identifique qué está realmente construyendo el candidato y qué está copiando del texto fuente (Cumming 2013, Cumming et al. 2005, Staples et al. 2013, Weigle 2002, Weigle y Parker 2012). Para poder llevar a cabo dicha tarea, se necesitaría de programas que permitan evidenciar el calco de paquetes léxicos. Quizás una solución a este problema sería ofrecerle lecturas al sustentante y permitirle que tome notas para que extraiga las ideas y argumentos principales enunciados en los textos. Posteriormente, retirarle los textos y asignarle las instrucciones sobre lo que tendrá que escribir con base en lo que previamente leyó; esto con el fin de que demuestre su capacidad de construir el conocimiento con base en las lecturas. No obstante, habría que emplear programas computacionales para poder evidenciar lo que se ha copiado literalmente en el texto.

En relación con la segunda hipótesis, en el caso de las gráficas, resulta evidente que la interpretación de estas no siempre es sencilla. Sin embargo, debido al creciente uso de estas en los artículos de investigación en las ciencias sociales, resulta pertinente incluirlas para ver de qué manera los candidatos describen y contrastan información (Hyland 2009, Shah y Freedman 2009). Aunado a lo anterior, la gran ventaja de estas radica en que el estímulo visual contiene los elementos necesarios para que los estudiantes escriban, con lo cual se evita que estos copien fragmentos de los textos fuente en tareas integradas. No obstante, la selección de gráfica no es una tarea fácil, pues es importante contemplar que esta desplegará cierto léxico que el examinado debe poseer, sin importar los estudios académicos que desee cursar o esté por cursar; razón por la cual estas no deberán contemplar léxico especializado (Yang 2012), dado que esto limitaría la capacidad del candidato para explotar al máximo el potencial de escritura y además coartaría la capacidad de escritura en aquellos candidatos con menor aptitud. Entonces, resulta importante seleccionar gráficas de barra o de pastel que sean sencillas y claras para los sustentantes (Uysal 2010). Finalmente, lo ideal sería que quienes seleccionan las gráficas, así como los evaluadores, se dieran a la tarea de responder la tarea, puesto que solamente así podrían identificar posibles problemas que los candidatos pudieran tener al momento de escribir y, consecuentemente, tomar acciones previas a la aplicación de la prueba.

En el caso de los ensayos de carácter argumentativo, como se pudo apreciar en la revisión bibliográfica, el contenido de estos permitirá que se establezca o no un lazo interactivo entre el sustentante y la tarea (Bachman 2002, Kormos 2011). Sin embargo, la selección de temas generales resulta pertinente para facilitar esta interacción en los sustentantes (Lee 2008, Lee y Anderson 2007, Hamp-Lyons 1993, He y Shi 2012, Weigle 2002). Por ello, más que abogar en favor o en contra de este tipo de tareas escritas, resulta importante incorporar temas que resulten interesantes para los candidatos, quienes claramente poseen un sentido crítico y están dispuestos a participar siempre y cuando los temas estimulen su capacidad de análisis. Consecuentemente, no resulta trivial que Petersen (2009) haya encontrado casos en los cuales los estudiantes se opusieron abiertamente a contestar este tipo de ensayos, dado que cuando estos perciben que su capacidad crítica se ve minimizada con temas absurdos o fuera de lugar, estos reaccionan negativamente.

En relación con el modo de transcripción, los hallazgos aquí encontrados contradicen la postura de otras investigaciones que afirman que este modo de transcripción demuestra mejores resultados en los sustentantes y también es preferido por estos (Horkay et al. 2006, Mogey et al. 2010, Mogey y Hartley 2013, Weigle 2002). Es muy evidente que la escritura a mano distorsiona la realidad del texto y favorece o demerita la apreciación subjetiva por parte del evaluador. Sin embargo, la costumbre de evaluar textos escritos a mano dificulta el cambio a aquellos escritos a computadora, dado que parecen ser más cortos y productos finales revisados y editados, cuando en realidad no es así. Aunado a lo anterior, la estrategia de corregir en el papel ha sido siempre vista como un aspecto positivo de la escritura en segundas lenguas, misma que se esfuma al no poder evidenciar lo que el sustentante ha editado en una computadora. Otra cuestión importante radica en que

los participantes que escribieron sus textos a computadora ya son alumnos de posgrado y, por lo tanto, están familiarizados con el teclado en español; sin embargo, si un candidato que apenas desea ingresar al ciclo universitario no estuviera familiarizado con este, posiblemente se vería en desventaja, dependiendo de la práctica que haya tenido previamente con el teclado en el idioma que desea escribir. Entonces, valdría la pena proponer una modalidad abierta; es decir, que el candidato escogiera la manera mediante la cual escribirá sus textos.

Retomando a Weigle (2002), pese a que las variables que se ven involucradas en las tareas de escritura no esclarecen el panorama de cuáles serían óptimas para evaluar en un contexto académico, estas no son insuperables. De hecho, entre mayor sea el conocimiento de qué variables inciden positiva o negativamente en las tareas de escritura, mayor será la capacidad de controlarlas o de generar tareas estandarizadas que permitan evitar sesgos en los resultados y, en consecuencia, garantizar una evaluación más justa y equitativa. De hecho, la inclusión de estas tareas escritas de respuesta construida está abriendo un amplio campo para la evaluación. Sin lugar a dudas, sería importante considerar la posibilidad de emplear tareas de respuesta construida, aunque para ello se requiere de recursos tecnológicos que den cuenta de lo que se ha calcado del texto, además de que la escritura dependería de la correcta selección de textos, tarea que incrementa los recursos humanos que se requieren para el diseño de este tipo de tareas. Por añadidura, en caso de que se desee trabajar con esta modalidad, será importante demostrar que estas en realidad permiten evaluar la habilidad de dar crédito adecuado a las fuentes y de integrar el material citado de forma adecuada.

Finalmente, una de las principales limitaciones del presente estudio se debe a que no se abordan las variables del candidato ni se cuenta con estudios que evidencien la interacción del sustentante con las tareas de escritura. Por ende, en futuras investigaciones sería importante conocer la opinión de los sustentantes respecto de la tarea propuesta; es decir, mediante entrevistas, protocolos de pensamiento en voz alta (*think-aloud protocols*) al momento de resolver la tarea y cuestionarios abiertos o cerrados, indagar sobre posibles problemas o dificultades que pudieran encontrar con la temática, las gráficas, la consigna de la tarea, la relevancia del tema a escribir, el tiempo otorgado y otros factores que pudieran incidir en la resolución de esta (espacio, mobiliario, orden de la tarea con respecto a las otras pruebas, etc.).

[1] Bachman y Palmer (2010: 59) definen una tarea de uso del lengua “como una actividad que involucra a los individuos en el uso de la lengua con el propósito de alcanzar una meta u objetivo en una situación particular”. Entonces, afirman los autores que las tareas deben estar claramente orientadas hacia un propósito, situadas en un contexto particular y que involucran la activa participación de los usuarios de la lengua.

2 Ibid., p.22 hacen una clara descripción de aquellos componentes que garantizan la utilidad de los exámenes. En el caso de la autenticidad, los autores la definen “como el grado de correspondencia de las características de una tarea de lengua en un examen con las características de una tarea real en la cual se usa la lengua”. Por lo tanto, los autores consideran que para poder diseñar una tarea que auténtica, es importante describir el dominio en el cual se hace uso de la lengua meta.

3 Lindqvist, Gudmundson y Bardel (2013: 111) definen la sofisticación léxica como el porcentaje de palabras difíciles empleadas dentro de un texto. El método que se emplea para llevar a cabo dicha tarea se conoce como *Lexical Frequency Profile* (LFP). El LFP mide la proporción de palabras de alta frecuencia contra la proporción de baja frecuencia en un texto. Todas las palabras son divididas de acuerdo a bandas previamente establecidas que han sido generadas mediante corpus escritos. Por ejemplo, el programa Vocabprofile establece la división entre las primeras 1000 palabras más frecuentes versus las siguientes 1000 palabras más frecuentes, además de una lista de palabras consideradas como académicas.

4 Los textos que a continuación se transcriben presentan algunos errores ortográficos y de acentuación, pues así fue como los escribieron los estudiantes.

5 Se destaca en negritas lo que aparecía en la consigna de las tareas escritas empleadas en el pilotaje.

REFERENCIAS BIBLIOGRÁFICAS

Alderson, J.C. y Clapham, C. (1993). *Examining the ELTS Test: An Account of the First Stage of the ELTS Revision Project* – Research Report 2. The British Council/University of Cambridge Local Examinations Syndicate.

Bachman, L. (2002). Some reflections on task-based language performance assessment. *Language Testing*, 19: 453-476.

Bachman, L., y Palmer, A. (1996). *Language Testing in Practice*. Oxford University Press.

Bachman, L. y Palmer, A. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Reino Unido: Oxford University Press.

- Canale, M. (1983). From Communicative Competence to Communicative Language Pedagogy. En J. Richards y R. Schmidt (Eds.). *Language and Communication*. Londres: Longman.
- Canale, M., y Swain, M. (1980). Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing. *Applied Linguistics*, 1 (1): 1-47.
- Chan, S. H., Wu, R. Y., Weir, C. J. (2014). *Examining the context and cognitive validity of the GEPT Advanced Writing Task 1: A comparison with real-life academic writing tasks*. LTTC-CRELLA Collaboration Project RG-03. Taiwan.
- Cumming, A. (2013). Assessing Integrated Writing Tasks for Academic Purposes: Promises and Perils. *Language Assessment Quarterly*, 10 (1): 1-8.
- Cumming, A., Kantor, R., Baba, K., Erdosy, U., Eouanzoui, K., James, M. (2005). Differences in written discourse in independent and integrated prototype tasks for Next Generation TOEFL. *Assessing Writing*, 10: 5–43.
- East, M. (2006). The impact of bilingual dictionaries on lexical sophistication and lexical accuracy in tests of L2 writing proficiency: A quantitative analysis. *Assessing Writing*, 11: 179–197.
- González, R. O. (Coord.) (2014). *Habilidades lingüísticas de los estudiantes de primer ingreso a las instituciones de educación superior del área metropolitana de la ciudad de México*, México: ANUIES
- Hamp-Lyons, L. (1993) Pre-text: Task-related influences on the writer. En Hamp-Lyons, L., editor, *Assessing second language writing in academic contexts*. Norwood, NJ: Ablex Publishing Corporation: 87–110.
- He, L. y Shi, L. (2012). Topical Knowledge and ESL writing, *Language Testing*, 29 (3): 443-464.
- Horkay, N., Bennett, R.E., Allen, N., Kaplan, B., Yan, F. (2006). Does it matter if I take my writing test on computer? An empirical study of mode effects in NAEP. *The Journal of Technology, Learning & Assessment*, 5: 4–49.
- Huot, B., O'Neill, P., Moore, C. (2010). A Usable Past for Writing Assessment. *National Council of Teachers of English*, 72 (5): 495-517.
- Hyland, T. A. (2009). Drawing a line in the sand: Identifying the border zone between self and other in EL1 and EL2 citation practices. *Assessing Writing*, 14 (1): 62-74.
- Kane, M. (2002b). Validating high-stakes testing programs. *Educational Measurement: Issues and Practice*, 21(1): 31–41.
- Kane, M. (2006). Validation. En R. Brennan, *Educational measurement*. 17–64. Westport: American Council on Education and Praeger.
- Kane, M. (2013). Validating the Interpretations and Uses of Test Scores, *Journal of Educational Measurement*, 50 (1): 1-73.
- Kane, M., Crooks, T., Cohen, A. (1999). Validating measures of performance. *Educational Measurement: Issues and Practice*, 18(2): 5–17.
- Kobrin, J. L., Deng, H., Shaw, E. J. (2011). The association between SAT prompt characteristics, response features, and essay scores. *Assessing Writing*, 16: 154–169.
- Kormos, J. (2012). The role of individual differences in L2 writing. *Journal of Second Language Writing*, 21: 390–403.
- Lindqvist, C., Gudmunson, A., Bardel, C. (2013). A new approach to measuring lexical sophistication in L2 oral production, en C. Bardel, C. Lindqvist y B. Laufer (Eds.). *L2 vocabulary acquisition, knowledge and use. New perspectives on assessment and corpus analysis*: 109-126. Eurosla Monographs Series, 2.
- Lee, H. (2008). The relationship between writers' perceptions and their performance on a field-specific writing test, *Assessing Writing*, 13: 93–110.
- Lee, H. y Anderson, C. (2007). Validity and topic generality of a writing performance test. *Language Testing*, 24: 307–330.
- Leki, I., Cumming, A., Silva, T. (2008). *A Synthesis of Research on Second Language Writing in English*. Routledge: New York.

- Mendoza, A. (en prensa). Las prácticas de evaluación docente y las habilidades de escritura requeridas a nivel posgrado. *Innovación Educativa*.
- Mogey, N. y Hartley, J. (2013). To write or to type? The effects of handwriting and word-processing on the written style of examination essays. *Innovations in Education and Teaching International*, 50 (1): 85-93.
- Mogey, N., Paterson, J., Burk, J., Purcell, M. (2010). Typing compared with handwriting for essay examinations at university: Letting the students choose. *ALT-J Research in Learning Technology*, 18 (1): 29-47.
- Moore, T. y Morton, J. (2005). Dimensions of difference: a comparison of university writing and IELTS writing. *Journal of English for Academic Purposes*, 4: 43-66.
- Petersen, J. (2009). "This test makes no freaking sense": Criticism, confusion, and frustration in timed writing. *Assessing Writing*, 14: 178-193.
- Plakans, L. (2010). Independent vs. integrated writing tasks: A comparison of task representation. *TESOL Quarterly*, 44 (1): 185-194.
- Plakans, L. y Gebriel, A. (2013). Using multiple texts in an integrated writing assessment: Source text use as a predictor of score. *Journal of Second Language Writing*, 22: 217-230.
- Shah, P. y Freedman, E. G. (2009). Bar and Line Graph Comprehension: An Interaction of Top-Down and Bottom-Up Processes. *Topics in Cognitive Science*: 1-19.
- Staples, S., Egbert, J. Biber, D., McClair, A. (2013). Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL iBT writing section. *Journal of English for Academic Purposes*, 12: 214-225.
- Toulmin, S. (1958). *The uses of argument*. Reino Unido: Cambridge University Press.
- Uysal, H. H. (2010). A critical review of the IELTS writing Test. *ELT Journal*, 64 (3): 314-320.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.
- Weigle, S. C. y Parker, K. (2012). Source text borrowing in an integrated reading/writing assessment. *Journal of Second Language Writing*, 21: 118-133.
- Yang, H. (2012). The language learning potential of form-focused feedback on writing. *English for Specific Purposes*, 31: 174-187.