

TRADUCCIÓN TRANSPARENTE: MÉTODOS CUANTITATIVOS Y CUALITATIVOS EN LA EVALUACIÓN DE LA TRADUCCIÓN

Bryan J Robinson Fryer
Departamento de Traducción e Interpretación
Universidad de Granada

RESUMEN

En la enseñanza de una asignatura tan sumamente práctica como es la traducción científico-técnica la calidad de la traducción lo es todo. Dicha calidad tiene que definirse en cuatro dimensiones: calidad de la traducción como proceso; calidad para el alumnado; calidad de la traducción como producto; y calidad para el profesor.

En el contexto de la Licenciatura en Traducción e Interpretación hemos realizado un análisis estadístico transversal del uso de un baremo de descriptores, previamente hecho público, diseñado para evaluar las traducciones partiendo de los resultados de dos muestras paralelas. Nuestros resultados indican el valor del instrumento para establecer objetivos claros y transparentes. Las distribuciones y una serie de correlaciones aportan una descripción del alumnado y del comportamiento del baremo como instrumento que a la vez afianzan su valor y apuntan a la necesidad de afinar el instrumento en sí. Concluimos que esta clase de evaluación transparente significa un avance importante en la calidad de la enseñanza universitaria.

Palabras clave: Evaluación; Baremo de descriptores; Traducción.

ABSTRACT

The single most important aspect of the teaching of an entirely practical subject like Translation is the quality of the translation. This must be defined from four points of view: the quality of the translation process; quality as perceived by the student; the quality of the translation as a product; quality as perceived by the teacher.

In the context of first degree studies in Translation and Interpreting we carried out a cross-sectional statistical analysis of the use of a set of published criterion descriptors designed specifically to evaluate translations. The study was based on two parallel samples. Our results indicate the value of the instrument as a means of establishing clear, transparent objectives. Frequency distributions and Pearson correlation coefficients provided us with a description of the student sample and of the functioning of the descriptors which support our belief in the value of the instrument itself, as well as indicating refinements.

We conclude that this type of transparent evaluation offers a significant step forward in improving the quality of university teaching.

Keywords: Evaluation; Criterion descriptors; Translation Studies.

INTRODUCCIÓN

En la enseñanza de las asignaturas de traducción, dentro de la nueva Licenciatura en Traducción e Interpretación cuya primera promoción en Granada se acaba de licenciar, la cuestión fundamental de la "calidad" de la traducción causa estragos en los planteamientos de ordenación docente porque las asignaturas de traducción - Traducción General, Traducción Especializada, etc - se contemplan como teórico-prácticas a pesar de que traducir es una

ocupación práctica. Las asignaturas teóricas son la Lingüística aplicada a la traducción, la Traductología, la Terminología, la Documentación, o la Informática aplicada. Éstas son asignaturas de estudio, de clases magistrales, de apuntes y “codos”. Al contrario, en las clases de traducción la aptitud de cada uno para traducir se pone a prueba constantemente y las herramientas con las que se mide esta aptitud muchas veces se pierden en frases como “no me suena bien”, o “a mí, me suena mejor así”. ¿Se consideran éstos criterios de calidad? ¿Se aceptarían en una Facultad de Ciencias? ¿Son válidos para la evaluación?

CRITERIOS DE EVALUACIÓN

La calidad en la traducción tiene que definirse en cuatro dimensiones: calidad de la traducción como proceso; calidad para el alumnado; calidad de la traducción como producto; y calidad para el profesor. Vamos a definir estas cuatro “calidades”.

Calidad de la traducción como proceso

La importancia de definir la traducción como proceso yace en el mismo contexto formativo de la enseñanza universitaria. Traducir es un proceso, y el alumnado adquiere conocimientos y dominio de este proceso a lo largo de sus estudios, por lo cual el profesorado debe calcular el progreso del alumnado y así identificar el dominio del proceso adquirido en un momento dado, para luego poder dirigirlo al alumnado en el proceso de ampliar este dominio. El progreso del traductor-aprendiz no es lineal, sino *staccato*. El alumnado adquiere conocimientos y consigue avances para luego encontrarse estancado en un punto u otro. Pero, ¿cómo sabe el alumno individual en dónde se ha quedado? Aquí hace falta un instrumento de medición que no sea tan rudo como el examen final, y que sirva para indicar qué elementos del proceso traductor domina cada uno, y cuales son los que necesita mejorar.

Calidad para el alumno

Esta calidad viene a ser “la calidad necesaria como para que este profesor me apruebe”. Es una definición clara y explícita, pero poco detallada, y en el transcurso de la revisión individual de un examen probablemente puede ampliarse, si es que en diez minutos, a los diez días de realizar el examen, el individuo puede realmente acordarse del texto original, de su traducción - “Pues sí, es mi letra” - y de la intencionalidad de su elección de tal verbo, o tal orden de palabras. Entre esta y nuestra primera definición existen pocas coincidencias. Son definiciones mínimas, aparentemente subjetivas, y que poco tienen que ver con las dos definiciones que nos faltan.

Calidad de la traducción como producto

Se supondría que esta calidad es la que se evalúa con facilidad. Evidentemente, a la hora de estudiar una traducción se compara el texto fuente (TF) con el texto término (TT) y, sobre todo cuando ninguno de los implicados ha realizado la traducción, se llega a una serie de conclusiones en cuanto a la “calidad” del trabajo. Desgraciadamente no es tan fácil. No es tan fácil entre profesores y alumnado, y tampoco lo es - me atrevo a decir - entre muchos de los usuarios de traducciones: o sea, los clientes. El grado de especialización del texto, el uso que se va a hacer del TT, la rapidez con la que hay que producir la traducción, y el importe de la traducción, son algunos de los muchos factores sobre los que habría que llegar a un acuerdo previo antes de entrar en el proceso de evaluación.

Calidad para el profesor

La calidad para el profesor puede definirse como el objetivo de la asignatura: ¿qué nivel mínimo de habilidad tendrá que demostrar el alumnado para aprobar la asignatura? Consciente o inconscientemente el profesor actúa con referencia a un baremo de calidad cuando corrige traducciones. Los criterios de este baremo se hacen explícitos a través de los comentarios, exclamaciones, cruces, tachones y rayas, en fin, la tinta roja que vierte en el texto que tiene que corregir. Los comentarios individuales o en clase también sirven como ampliación de estos criterios. En la Universidad de Granada - como en todas las universidades - las “idiosincrasias” del profesorado - incluidas las mías - dan lugar al comentario frecuente del alumnado al efecto de que “Ah, sí, a Vd. le gusta que se edite mucho el texto”, por ejemplo. Nuestra definición global de calidad, pues, debería de contemplar la reconciliación de esta dimensión con las tres anteriores.

LA TEORÍA “CLÁSICA” DE EVALUACIÓN (Bachman 1990; Alderson et al 1995)

Los dos conceptos fundamentales en la evaluación son la validez y la fiabilidad del instrumento.

Validez

La validez se define como el grado en que una prueba mide las habilidades, destrezas, etc, que pretende medir. Tiene dos facetas internas: la validez aparente del instrumento para el consumidor, en nuestro caso el alumnado - y la validez con respecto al contenido que se supone que el profesor ha impartido anteriormente. También tiene una validez externa: la validez concurrente entre distintas convocatorias.

Fiabilidad

La fiabilidad del instrumento, cuando se trata de un baremo de descriptores, es la relación que se describe entre la aplicación del mismo baremo por más de un profesor: ¿utilizan la misma herramienta de la misma manera? También existe la fiabilidad interna del profesor individual: evidentemente, si yo no utilizo el baremo de descriptores de forma consistente, difícilmente podremos encontrar un punto de fiabilidad en la comparación de los resultados de los exámenes que he corregido con los que haya corregido cualquier compañero.

La corrección de un examen tiene obligatoriamente que ser válida y fiable para asegurar la calidad de la evaluación. Mathews (1985:90) señaló tres fuentes de inconsistencia en la aplicación de un examen, que podrían derivarse de la especificación original del examen; de la producción y/o moderación del examen como instrumento; o de la corrección del examen, en cuanto a variedad que puede existir en la manera en la que individuos interpreten y/o apliquen un baremo de corrección. El objetivo ideal es la producción de un instrumento de corrección estándar, que luego puedan utilizar individuos que hayan recibido la formación suficiente como para aplicarlo de manera consistente.

INSTRUMENTO: EL BAREMO DE DESCRIPTORES

Definimos un baremo de descriptores según Alderson *et al* (1995:287-88). Consiste en unas “etiquetas”, p.ej. “A”, “B”, “C”, “1”, “2”, “3”, “Sobresaliente”, “Notable”, “Aprobado”, para

cada una de las cuales existe una afirmación del comportamiento que cada punto en la escala significa. Nuestra escala puede ser de intervalos: por lo cual entenderíamos que la diferencia entre un “6” y un “7”, sería igual a la diferencia entre un “7” y un “8”. O puede ser ordinal: por lo cual la diferencia entre un “Aprobado” y un “Notable” no necesariamente sería igual que la diferencia entre un “Notable” y un “Sobresaliente”. Y tampoco sería la diferencia entre un “Aprobado” y un “Sobresaliente”, dos veces la diferencia entre el “Notable” y el “Sobresaliente”.

Calificación	Diferencia	Calificación	Diferencia	Calificación	Diferencia
“1”		“A”	+/- ?	“Sobresaliente”	+/- ?
“2”	+ 1	“B”	+/- ?	“Notable”	+/- ?
“3”	+ 1	“C”	+/- ?	“Aprobado”	+/- ?

Pueden parecer detalles de poca importancia, pero realmente son importantes a la hora de desmitificar mucho del proceso de la corrección de exámenes con el fin de asegurar una mejora en la calidad de la enseñanza universitaria.

Los baremos de descriptores suelen ser o bien holísticos o bien analíticos. Un baremo holístico (Apéndice A) es, aparentemente más sencillo que uno analítico, y de hecho tiende a reducir el ámbito de discrecionalidad en la corrección. Pero la sencillez se pierde cuando uno comienza a leerlo con detenimiento. Un baremo analítico muestra de forma clara y transparente mucho de lo que el baremo holístico intenta. Nuestro instrumento (Apéndice B) es un baremo analítico que define cuatro criterios, y los subdivide en cinco niveles enumerados de mejor (5) a peor (1): suman 20 puntos máximo. Una característica importante es el hecho de que no aparece ninguna descripción que puede indicar en que consiste un “Sobresaliente” o cualquier otra calificación.

Objetivos

Los objetivos de nuestro estudio son:

- Describir la validez de nuestro baremo como instrumento de evaluación.
- Probar la fiabilidad del baremo.
- Indicar la manera en la que el baremo puede, también servir como herramienta didáctica.

En este estudio queremos evaluar la validez del instrumento que hemos diseñado desde el punto de vista externo, mediante el estudio de una serie de correlaciones entre los resultados de alumnos en distintas pruebas y en distintas épocas.

Hipótesis

La validez interna para el alumnado se apreciará en la manera en la se entiende y acepta el instrumento. También se verá en la apreciación por parte de un experto en la materia ajena a la investigación de los resultados. La validez externa se verá en la comparación entre los resultados que genera el uso de baremo en distintas pruebas.

La fiabilidad interna del instrumento se verá en las distribuciones de frecuencias producidas por las notas en distintas convocatorias.

Método

Este estudio pseudo-experimental tuvo como objetivo evaluar el uso de un baremo de descriptores de evaluación analítico. El baremo que nosotros presentamos (Robinson 1996: Appendix B) se preparó según las pautas que hemos descrito anteriormente (Robinson 1994). Los fundamentos para su uso variaron con respecto a nuestras intenciones en la medida en que se introdujo el baremo como herramienta activa en la enseñanza de la asignatura en el segundo año y se modificó el plan de evaluación para estos alumnos a raíz de los resultados ya obtenidos.

Contexto

Hemos recogido las notas del alumnado que siguió el bloque de Traducción científico-técnico de la asignatura Traducción especializada I A-B (Inglés) durante los cursos académicos 1995-96 (n=>125) y 1996-97 (n=>113). Este bloque es el equivalente a cinco créditos, es decir 50 horas lectivas impartidas a lo largo de 12-13 semanas.

Instrumento

El formato del examen - típico en la formación de traductores e intérpretes - consistió en la presentación de una "hoja de encargo" (Apéndice C) al menos quince días antes de la fecha del examen en la que se informaba al alumnado acerca de:

- El tema del texto fuente (1995-96) o las palabras clave (1996-97).
- La longitud aproximada del TF.
- Los parámetros tipológicos del TT: lugar de publicación (p. ej. revista, prensa); sección (p. ej. Ciencia, Internacional); grado de especialización (p. ej. no especializado o semi-especializado); supuesta fecha de publicación del TT.

En las asignaturas de traducción es una práctica aceptada permitir al alumnado utilizar libros de referencia además de cualquier recurso documental que haya preparado: glosarios, textos paralelos, etc.

MUESTRAS

La muestra "A" corresponde a los alumnos del curso 1995-96, y la "B" a los del 1996-97. Las diferencias más significativas entre las dos muestras fueron que:

- A la muestra "A" se le impartió el bloque de Traducción científico-técnica en el segundo cuatrimestre, cuando el alumnado había terminado el bloque económico-jurídico. En 1996-97, se impartieron los bloques al revés. Es importante señalar este hecho, ya que un cuatrimestre de clases prácticas puede suponer avances grandes en el alumnado que se tendrían que observar en las notas. Sería, pues, previsible que el rango de las notas de la muestra "A" fuera más alto que el de la muestra "B".
- La muestra "A" fue la primera promoción de la licenciatura, por lo cual, entre otras características especiales, no había "repetidores". En la muestra "B" se incluyeron 5 (4%)

“repetidores”. Un porcentaje tan reducido no debería de influir en los parámetros de la muestra.

- La muestra “A” presentó tres ejercicios evaluativos: un trabajo realizado fuera de clase y entregado en la 5ª semana de instrucción; un examen parcial durante las horas de clase después de 6 semanas; un examen final. A consecuencia de los resultados obtenidos en el primer año del estudio la muestra “B” no realizó más que el examen final.

Validez interna

Con el fin de averiguar la validez interna para el consumidor del baremo para la muestra “B” realizamos una importante innovación en la enseñanza de la asignatura - antes imposible debido a la novedad de la licenciatura en Traducción e Interpretación - que fue el uso y el debate abierto sobre el baremo durante las clases. Debido a que ya existía un número de textos calificados en el curso anterior se procedió a presentar una selección de éstos al alumnado con la finalidad de que ellos aprendieran a apreciar los criterios de evaluación y a entender como éstos podrían aplicarse para así poder ellos mismos evaluar sus propias traducciones. La clase se desarrolló de la siguiente manera:

- Se repartieron una serie de traducciones - pero sin que se pudiera identificar el autor - de un texto que el alumnado acababa de traducir.
- En equipos (max=5), el alumnado debatió el orden de calidad de las traducciones, anotando la justificación de sus decisiones.
- Los grupos cambiaron, e individualmente los alumnos compararon sus decisiones y debatieron las justificaciones.
- Se repartieron copias del baremo entre los grupos originales y volvieron a ordenar las traducciones ya contrastando y ordenando sus propios criterios de evaluación con respecto a los allí expuestos.
- Se discutieron los descriptores y se propusieron mejoras.

La validez interna con respecto al contenido de la asignatura se analizó a raíz del juicio de una compañera¹, profesora del otro bloque de la asignatura, de la siguiente manera:

- Se le entregó una muestra (10%) de exámenes representativos de las fronteras entre las calificaciones ordinales “Suspendido”, “Aprobado”, “Notable” y “Sobresaliente”. Definimos “frontera” como lo que se entiende por un “Suspendido alto”, un “Aprobado bajo”, etc. Ninguno de los textos tenía señalado la nota numérica u ordinal, y se había borrado la “tinta roja”.
- Se le pidió ordenar las traducciones en orden de calidad de mejor a peor, y apartar los “Suspendidos”.
- Después intercambiamos opiniones sobre los criterios de calidad.

De hecho no hubo acuerdo perfecto en cuanto al orden de calidad de entre las traducciones.

¹ Se agradece la colaboración de la profesora Catherine Way, Dpto de Traducción e Interpretación, Universidad de Granada.

No obstante, sí encontramos que hubo acuerdo en cuanto a la frontera entre el “Suspenso” y el “Aprobado”, y también lo hubo, *grosso modo*, en cuanto a las otras fronteras entre calificaciones. Como ambos conocíamos al alumnado este hecho parecía ayudarnos a la hora de juzgar los exámenes de calificaciones más “fronterizas”.

Validez externa

La validez del instrumento entre distintas convocatorias se estudió en el análisis estadístico de los tres ejercicios que realizaron la muestra “A”. En contra de lo previsto el rango de las notas numéricas era ligeramente más alto en el caso del examen final, y no del trabajo a entregar. La coincidencia entre los rangos del trabajo y del examen parcial tampoco era de esperar aunque probablemente se debía a la proximidad de las fechas en las que se realizaron los ejercicios.

	Trabajo	Examen parcial	Examen final
Rango	8 a 19	8 a 19	9 a 20

Las medidas de centricidad - media, mediana y moda - fueron muy similares entre los tres ejercicios, con dos diferencias en el caso del trabajo a entregar: las cifras fueron ligeramente más altas, y la curtosis fue negativa. Estos datos se pueden interpretar como que la mayor parte del alumnado sacó mejores notas en un trabajo realizado fuera de clase que en los exámenes: nada inusual.

	Trabajo	Examen parcial	Examen final
Media	13.85	13.1	13.37
Mediana	14	13	13
Moda	15	12	12
Curtosis	-0.97	4.68	0.3

Las coeficientes de correlación Pearson entre los tres ejercicios fueron el motivo de la supresión de dos de los tres ejercicios para la muestra “B” ya que fueron tan altas ($p = 0.001$):

	Trabajo	Examen parcial	Examen final
Trabajo	1		
Examen parcial	0.71	1	
Examen final	0.56	0.48	1

En el caso de la muestra “B” la distribución de frecuencias tenía características similares, pero las cifras fueron inferiores a las de la muestra “A”:

Examen final	Muestra "A"	Muestra "B"
Media	13.37	12.12
Mediana	13	12
Moda	12	11
Curtosis	0.3	2.96
Rango	9 a 20	7 a 18

¿Pueden ser estas diferencias simplemente el resultado de la inversión del orden en el que se impartieron los bloques? o ¿es que se ha aplicado el baremo de forma más rígida para la muestra "B"? o ¿significa que el alumnado de la muestra "A" es más competente que el de la muestra "B"? Nuestra respuesta, tras contrastar estos resultados con los del segundo bloque y los de la asignatura del curso anterior - Traducción General A-B - es que, efectivamente el nivel de competencia traductora de la muestra "B" es inferior al de la muestra "A": no son tan buenos traductores.

CONCLUSIONES

Estamos satisfechos con el diseño fundamental del baremo de descriptores y tenemos la intención de utilizarlo en el curso 1997-98 cuando éste se hará público con el programa de la asignatura. A raíz de nuestra experiencia y de las propuestas del alumnado es nuestra intención introducir algunas modificaciones en el texto de los descriptores individuales y añadir el nivel "0" a cada uno de los criterios. También pensamos hacer uso del baremo como herramienta de instrucción con más frecuencia en la revisión de las traducciones del mismo alumnado y ajenas. Nuestro próximo logro sería la extensión de esta metodología a otros miembros del departamento y otras combinaciones lingüísticas - Español-Alemán, Español-Francés - y a la otra especialidad: Traducción económico-jurídica para poder comprobar o no la validez y la fiabilidad del baremo en manos de otra.

APÉNDICE A

	DESCODIFICACIÓN	CODIFICACIÓN		
	<i>Contenido</i>	<i>Registro, vocabulario, términos.</i>	<i>Discurso del texto con respecto a los parámetros del TT.</i>	<i>Gramática, acentos, ortografía.</i>
5	Dominio completo del tema. La transferencia del contenido del texto se ha realizado con una precisión absoluta.	Expresión perfectamente adecuada a la intencionalidad, al tema y la función del TT.	El TT evidencia todas las características de un texto original en la LT. Organización del discurso a nivel de texto, párrafo, y oración completamente apropiada.	Impecable en todos los aspectos.

4	Buen dominio del tema. En unas pocas ocasiones se encuentran excesos o faltas de contenido. A veces la falta de claridad en el TT puede derivarse de dificultades de comprensión.	Léxico adecuado, aunque en alguna que otra ocasión se podría haber realizado una mejor elección. Registro completamente adecuado.	TT no perfecto, pero demuestra muchas características de un texto original en la LT. Buena organización del discurso tanto a nivel de oración como al de coherencia global.	Aparecen unos pocos errores de menor importancia.
3	Dominio suficiente. No aparece ningún error ni omisión de importancia, aunque alguno que otro de menor importancia. No comunica ningún dato erróneo.	Léxico aceptable por lo general. Aparecen algunos términos inapropiados, pero nunca impiden la lectura fluida del TT.	Organización del texto adecuada, aunque a nivel de oración aparecen estructuras forzadas o poco naturales.	Correcto en todos los aspectos gramaticales y de precisión en la presentación de datos. Aparecen unos pocos errores de ortografía y/o de acentos. También puede haber algún que otro error de puntuación.
2	Conocimiento parcial del tema. Aparecen errores de mayor y de menor importancia. Hay un número de omisiones de puntos de contenido importantes.	El uso de términos inapropiados impiden la comunicación. Registro inapropiado y/o inconsistente.	Elementos de incoherencia. En el texto aparecen estructuras calcadas del TO.	Hay errores gramaticales y/o de precisión en la presentación de datos. Aparecen bastantes errores de ortografía y/o de acentos.
1	Comprensión del contenido insuficiente. Comunica datos erróneos. Existen errores de mayor y de menor importancia. Se omite una cantidad importante del contenido.	Uso de léxico inapropiado e incorrecto. Poca evidencia del uso del concepto de registro.	Poca coherencia. El TT es una traducción formal en cuanto a la transferencia de estructuras gramaticales de la LO.	El TT es impresentable debido al número y/o la clase de errores de índole técnica que aparecen en ello.

APÉNDICE B

9 - 10	Capta todos los matices de la pregunta y responde con originalidad y claridad tanto de ideas como de estructura y expresión. Domina los conceptos. Maneja con soltura el registro propio del ensayo académico. Los ejemplos y las referencias bibliográficas son pertinentes y están presentados correctamente.
7 - 8	Demuestra haber captado el alcance general de la pregunta y responde con bastante acierto y claridad de argumentación. Su dominio de los conceptos es amplio aunque no completo. A veces no es coherente en el uso de las convenciones del género. Utiliza bastantes ejemplos y referencias bibliográficas aunque no siempre cita sus fuentes adecuadamente
5 - 6	Parece haber entendido el alcance general de la pregunta y la respuesta es adecuada aunque un tanto limitada. A veces la argumentación es confusa. Se perciben algunas lagunas en su dominio del tema. Su uso de las normas que caracterizan el ensayo académico no es siempre correcto. Cita algunos ejemplos y/o referencias bibliográficas aunque no siempre de forma apropiada. Hay algunos "lapsus" de ortografía y/o puntuación que no afectan a la comprensión del trabajo pero sí le restan calidad.

I Jornadas Andaluzas

3 - 4	No ha demostrado haber entendido el alcance global de la pregunta. La respuesta es inadecuada y a veces confusa. Demuestra ignorancia en aspectos importantes del tema. Demuestra una falta de conocimiento del género. Apenas utiliza ejemplos y/o referencias bibliográficas. Tampoco son siempre acertados los que utiliza. Existen errores frecuentes de ortografía y/o puntuación.
1 - 2	No ha entendido la pregunta. No demuestra conocimiento de los conceptos más elementales del tema. No conoce el género de texto. No utiliza ejemplos ni referencias bibliográficas. El texto está plagado de errores lingüístico-técnicos

APÉNDICE C

Hoja de encargo

Convocatoria ordinaria de exámenes junio-julio 1997

Examen cuatrimestral (parcial)

3º de LICENCIATURA

TRADUCCIÓN ESPECIALIZADA A-B (Inglés)

Bloque científico-técnica

Provisional date 16 June

Time 9.00-11.00

Room 17

Source text

Keywords: Multiple sclerosis; vaccine; Central nervous system

Newspaper feature on health

Parts of a longer text

200-250 words

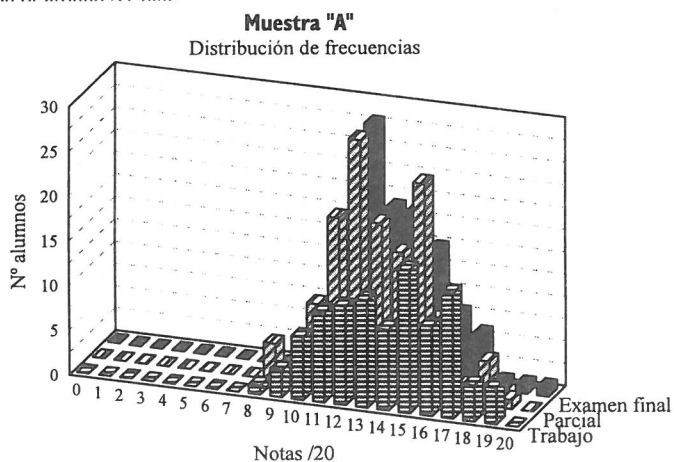
Target text

New Scientist magazine

'This Week' section

Complete article in itself

Assume publication on or around ST date



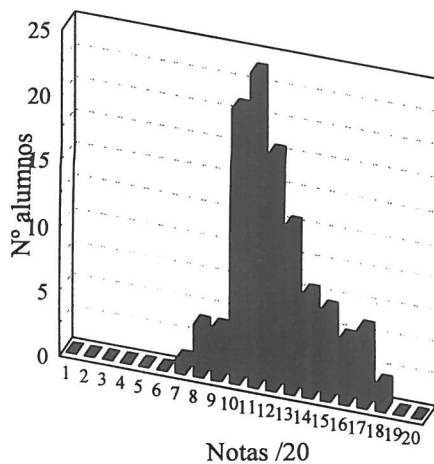
Estadística descriptiva

	Trabajo	Parcial	Examen final
Media	13,85	13,10	13,37
Error típico	0,26	0,23	0,20
Mediana	14,00	13,00	13,00
Moda	15,00	12,00	12,00
Desviación típica	2,76	2,54	2,11
Varianza	7,61	6,45	4,46
Curtosis	-0,97	4,68	0,30
Asimetría	-0,03	-0,90	0,38
Recorrido	11,00	19,00	11,00
Mínimo	8,00	0,00	9,00
Máximo	19,00	19,00	20,00
Suma	1.551,00	1.638,00	1.524,00
Categoría	112,00	12500	114,00
Nivel de confianza(0.950000)	51	45	39

Coefficientes de correlación Pearson

	Trabajo	Parcial	Examen final
Trabajo	1,00		
Parcial	0,71	1,00	
Examen final	0,56	0,48	1,00

Muestra "B"
Distribución de frecuencias



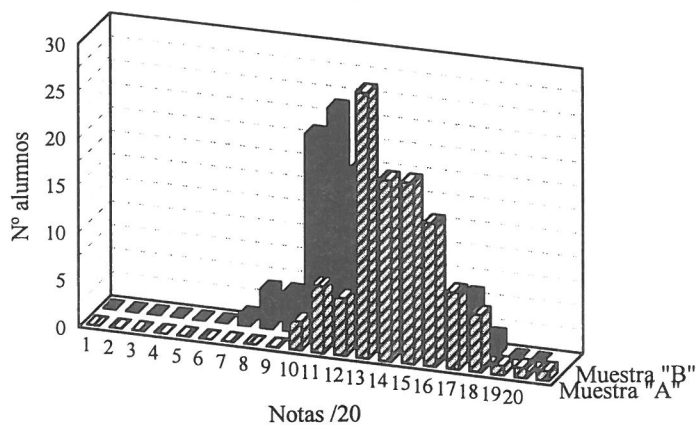
AP
ÉN
DI
CE
E

EXAMEN FINAL

MEDIA	12,12
ERROR ESTÁNDAR	0,22
MEDIANA	12,00
MODO	11,00
DESVIACIÓN ESTÁNDAR	2,39
VARIANZA	5,73
KURTOSIS	2,96
TERGIVERSADO	-0,32
RANGO	18,00
MÍNIMO	7,00
MÁXIMO	18,00
SUMA	1.369,00
CUENTA	113,00
NIVEL DE CONFIANZA(0,950000)	44

APÉNDICE F

Comparación "A" v. "B" - Examen final
Distribución de frecuencias



BIBLIOGRAFÍA

ALDERSON, J.C.; CLAPHAM, C. y WALL, D. (1995). *Language test construction and Evaluation*. Cambridge: CUP.
BACHMAN, L. (1990). *Fundamental Considerations in Language Testing*. Oxford. OUP.
HATIM, B. & MASON, I. (1990). *Discourse and the translator*. Harlow, Essex. Longman Group Limited.

- HOUSE, Juliane (1977). A Model for Assessing Translation Quality. *Meta*, Vol. XXII, N° 2, 103-109.
- ____ (1981). *A Model for Translation Quality Assessment*. Tübingen. Gunter Narr Verlag.
- INSTITUTE OF LINGUISTS (1995). *Marking guidelines for the Diploma in Translation*. London, Institute of Linguists. Documento interno, confidencial.
- INTERNATIONAL BACCALAUREATE (1993). *Language B Pilot Project Guide*. Cardiff. International Baccalaureate.
- MATHEWS, J.C. (1985). *Examinations: A Commentary*. London: George Allen and Unwin.
- ROBINSON, B.J. (1996). La evaluación de la traducción. *Actas de las I Jornadas sobre Diseño Curricular del Traductor e Intérprete*. Granada: F.T.I., Universidad de Granada, 1996:199-208.
- ____ (1994). Lightening the essay-marking load. *Greta*, Vol.1 N° 2, 19-24.
- SPOLSKY, B. (1995). *Measured Words*. Oxford: OUP.