



Instituto Nacional para la
Evaluación de la Educación

La Traducción de Pruebas en Comparaciones Internacionales: Un Estudio Preliminar

ASPECTOS LINGÜÍSTICOS DE LA
TRADUCCIÓN DE TIMSS-1995

Guillermo Solano-Flores
American Institutes for Research

Eduardo Backhoff Escudero
Universidad Autónoma de Baja California

LA TRADUCCIÓN DE PRUEBAS EN COMPARACIONES INTERNACIONALES: UN ESTUDIO PRELIMINAR

Guillermo Solano-Flores * y Eduardo Backhoff Escudero**

*American Institutes for Research **Universidad Autónoma de Baja California

Introducción

La noción básica de validez establece que el puntaje obtenido por un examinado en un instrumento no debe estar determinado por factores distintos ajenos a lo que se pretende medir (Messick, 1995). Un factor importante que influye en la validez de pruebas de desempeño académico es el lenguaje (American Educational Research Association, 1999). El lenguaje es necesario para poder aplicar una prueba, aun cuando el contenido que se evalúa sea, por ejemplo, la habilidad matemática o los conocimientos de ciencias naturales, no la habilidad verbal la lectura o la escritura.

En años recientes se ha investigado la influencia del lenguaje en el desempeño de los alumnos en pruebas. Se sabe que la manera en que un estudiante interpreta un reactivo es muy sensible a la forma en que está redactado. Por ejemplo, una sola palabra puede hacer una enorme diferencia en lo que los estudiantes creen que tienen que hacer para resolver un problema (e.g., Baxter, Shavelson, Goldman, y Pine, 1992). Además, la combinación de las experiencias personales de los estudiantes y la manera en que se presenta la información de un reactivo puede determinar que el alumno lo interprete correcta o incorrectamente (Solano-Flores y Trumbull, 2003; Solano-Flores, Trumbull y Nelson-Barber, 2001).

Como resultado de la globalización de la economía, se presenta ahora con más frecuencia la necesidad de traducir instrumentos a idiomas en los que no fueron escritos originalmente o de adaptarlos para poblaciones para los que no fueron creados (Hambleton, 1994). Las comparaciones internacionales han planteado retos importantes para el desarrollo de instrumentos en varios idiomas y han hecho necesario perfeccionar los procedimientos para la adaptación lingüística de pruebas (Grisay, 1998).

Este informe presenta los resultados de un estudio exploratorio de los aspectos lingüísticos asociados a la traducción de las pruebas el Tercer Estudio Internacional de Matemáticas y Ciencias (TIMSS, por sus siglas en inglés). Aunque México retiró su participación de la comparación internacional antes de que se publicaran los resultados, se conservaron copias de las pruebas con las que se evaluó a los alumnos mexicanos. La comparación de estas pruebas con los originales en inglés de los que se les tradujo permite hacer un estudio cuidadoso de la calidad de la traducción.

Se plantearon dos objetivos: 1) efectuar una evaluación preliminar de la calidad de la traducción de los reactivos TIMSS-95, versión mexicana; y 2) desarrollar y poner a prueba la versión inicial de un marco conceptual y metodológico para un estudio completo de la calidad de la traducción de los reactivos TIMSS-95 México. El objetivo a largo plazo de investigaciones de esta naturaleza es poder ofrecer un conjunto de lineamientos metodológicos para la traducción de pruebas en proyectos internacionales en los que México participe en el futuro. Dichos lineamientos metodológicos estarán orientados a la reducción del sesgo al comparar poblaciones diferentes.

Tradicionalmente, la evaluación del sesgo de reactivos se basa en comparar grupos poblacionales que difieren en algún factor como el sexo, el nivel socioeconómico o la lengua materna (por ejemplo, Camilli, y Shepard, 1994; Van de Vijver y Tanzer, 1998). A diferencia de esos enfoques, este estudio se ocupó de los aspectos formales de los reactivos TIMSS relacionados con el

idioma: las propiedades de los reactivos que pueden ser relevantes desde un punto de vista lingüístico y que son evaluadas por jueces de acuerdo con un conjunto preestablecido de criterios de corrección de traducción o redacción de reactivos (véase Hambleton y Jones, 1994; y Solano-Flores, 2003).

Antecedentes

En 1995 se llevó a cabo el Tercer Estudio Internacional de Matemáticas y Ciencias (TIMSS, por sus siglas en inglés). En esta comparación internacional que involucró a 41 países se aplicaron pruebas a dos muestras de estudiantes: Población 1, constituida por alumnos de 3º y 4º grado, y Población 2, constituida por estudiantes de 1º y 2º grados de secundaria. Estas pruebas consistieron en un conjunto de reactivos de opción múltiple y de respuesta abierta de ciencias naturales y de matemáticas.

Uno de los principales retos planteados por TIMSS deriva de aplicar el mismo conjunto de reactivos a poblaciones de distintas culturas que hablan distintos idiomas. A fin de que los resultados obtenidos no estuvieran influidos por diferencias extremas en el contenido curricular, representantes de todos los países participantes en TIMSS estuvieron involucrados en el proceso de decidir cuáles reactivos debieran ser incluidos en el instrumento. A fin de reducir el efecto del idioma o las diferencias entre dialectos de un mismo idioma como fuente de invalidez del instrumento, cada país se hizo cargo de la traducción de los reactivos, originalmente escritos en inglés.

Aunque México participó en este proyecto internacional, el gobierno mexicano solicitó su retiro poco después de que los datos de la comparación se dieran a conocer entre los países participantes, pero antes de que se les publicara. Los acuerdos de TIMSS concedían a los países la prerrogativa de retirarse de la comparación internacional en cualquier momento, en cuyo caso todos los datos de ese país debían ser destruidos de las bases de datos de TIMSS y ser excluidos de todos los análisis. Al retirarse México de TIMSS, se volvieron a efectuar todos los análisis de datos como si nunca hubiera participado en el estudio.

Afortunadamente, se conservaron copias de todos los cuadernillos que se les dieron a los estudiantes, lo que permite efectuar un análisis formal detallado de la equivalencia lingüística de los reactivos y de algunos otros factores asociados a su traducción. El formato de los cuadernillos y la secuencia de los reactivos en español son exactamente los mismos que los de los materiales en inglés. Los cuadernillos tienen tanto reactivos de opción múltiple como reactivos de respuesta abierta.

Traducción de pruebas TIMSS-1995 y revisión de las traducciones

Desarrollar pruebas es un proceso complejo y laborioso que requiere el refinamiento continuo de tres componentes: el planteamiento de un problema o tarea, un formato de respuesta (como el cuadernillo en el que se capturan las respuestas de los estudiantes) y un sistema de calificación (Solano-Flores y Shavelson, 1997). La mayor parte de este continuo perfeccionamiento tiene que ver con la manera de plantear un problema, las palabras que se usan y la forma en que los estudiantes interpretan los reactivos y responden a esos reactivos. Al final de ese proceso se alcanza un equilibrio delicado: cualquier modificación a la redacción del reactivo (por ejemplo, una palabra o el orden de una oración) puede producir formas diferentes, no anticipadas ni deseadas, en las que los estudiantes pueden interpretar el reactivo. Este es el estado de desarrollo de una prueba en que se considera que se ha alcanzado una eficiencia máxima en el uso del lenguaje.

Cuando una prueba es traducida a otro idioma, es difícil que tenga lugar el mismo proceso de refinamiento que ocurrió cuando se le escribió en el idioma original. Mientras que el desarrollo de la prueba en la lengua original llega a tomar muchos meses, generalmente la traducción del instrumento se lleva a cabo en tiempo muy breve, en el mejor de los casos con una sola iteración

de revisiones (Valdés y Figueroa, 1994). Si no se efectúa con el cuidado suficiente, la traducción puede llegar a constituir una fuente de sesgo de método (véase van de van de Vijver y Portinga, 1997) que atenta contra la validez del instrumento.

El documento “Traducción de pruebas de rendimiento académico utilizadas en estudios internacionales” (Hambleton, 1992), constituyó la base para los lineamientos para la traducción de pruebas que se proporcionaron a los países participantes en el Estudio TIMSS-1995. Sin embargo, cada país fue responsable de implementar la manera de aplicar esos lineamientos. Más aún, debido a limitaciones económicas y de tiempo, no todos los países participantes aplicaron estos procedimientos (Maxwell, 1996). Ello hizo que la calidad de las traducciones dependiera más de los procedimientos de revisión de las pruebas (a cargo de una compañía de traducción ubicada en Canadá) y de la competencia misma de los traductores empleados en cada país.

Preguntas de investigación

En este estudio se hizo una evaluación preliminar de la calidad de la traducción de los reactivos TIMSS-1995. Se plantearon dos preguntas de investigación:

- 1) ¿Cuáles son los errores que limitan la calidad de la traducción y pueden afectar la validez de los reactivos TIMSS-1995-México?
- 2) ¿Cómo varía el tipo o la frecuencia de esos errores según el área de contenido (ciencias naturales y matemáticas) y según la población (Población 1 y Población 2).

Métodos

Para contestar tales preguntas, se elaboró un marco conceptual para el análisis de los reactivos, se desarrolló un sistema de codificación de errores, y se aplicó ese sistema de codificación para analizar una muestra de los reactivos TIMSS-1995-México.

Marco conceptual

Un conjunto creciente de investigaciones sobre idioma, lenguaje y pruebas muestra que nunca debe subestimarse la importancia de ningún aspecto involucrado en la traducción de reactivos. Se sabe, por ejemplo, que la manera en que un reactivo presenta la información puede afectar la manera en que los estudiantes interpretan ese reactivo (Solano-Flores y Nelson-Barber, 2001). Muchas traducciones de reactivos que pueden ser consideradas como correctas pueden no reflejar el estilo discursivo natural del idioma objetivo, lo que puede implicar, entre otras cosas, que se incremente la demanda cognoscitiva del reactivo (Solano-Flores, Trumbull y Nelson-Barber, 2003). Más aún, a partir de investigaciones efectuadas con reactivos empleados en TIMSS-1995, se sabe que la simple sustitución de una sola palabra por otra en la traducción de una prueba es suficiente para modificar el funcionamiento diferencial de un reactivo (Ercikan, 1998).

Como parte básica del marco conceptual, en este estudio se desarrolló una tipología de los errores de traducción de reactivos. El desarrollo de esta tipología se basó en: 1) el sistema de codificación utilizado por los equipos de revisores de TIMSS-1995 (Mullis, Kelly y Haley, 1996) y 2) el marco conceptual para calificar errores de traducción que emplea la ATA como conjunto de criterios para calificar muestras de traducción de candidatos para la certificación de traductores (American Translators Association, 2003).

Los errores de traducción fueron agrupados en las siguientes siete categorías. Una definición detallada de ellas aparece en el Apéndice I de este informe.

- **Estilo y ortografía.** La traducción del reactivo no es consistente con las convenciones estilísticas y ortográficas utilizadas en textos impresos en el país.

- **Formato.** La información visual del reactivo en español se presenta de manera diferente al reactivo en el idioma original. Esta información visual incluye tablas, gráficas e ilustraciones.
- **Convenciones.** La traducción no refleja el estilo convencional de reactivos en el idioma español o viola principios de redacción de reactivos aceptados universalmente.
- **Gramática y sintaxis.** La estructura gramatical del reactivo no es tan natural como si éste hubiera sido escrito originalmente en español o no es adecuada para la edad del estudiante examinado.
- **Información.** La traducción altera la cantidad o el contenido de la información que se proporciona al estudiante para entender las instrucciones o el problema que se le presenta.
- **Semántica.** Las ideas expresadas en el reactivo no tienen el mismo significado que en la lengua original debido a que la traducción es imprecisa.

En general, los errores en las categorías 1, 2 y 3 pueden ser considerados como de menor gravedad que los errores de contenido (aunque ciertamente inaceptables en cualquier proyecto que involucra la producción masiva de materiales impresos). Sin embargo, no se debe subestimar el posible impacto de estos errores en la forma en que el estudiante puede interpretar un reactivo. La diferencia entre *como* y *cómo* o entre *casa* y *caza* son ejemplos de casos en que un acento o una letra puede llegar a determinar la forma en que se entiende un problema. La desviación en la versión en español de un reactivo en la forma en que aparece una gráfica puede también alterar el tipo de información que se proporciona al estudiante. Finalmente, la desviación de las convenciones en la redacción de reactivos puede provocar que al estudiante le resulte más difícil leer un reactivo cuando éste no aparece en el estilo estándar al que está acostumbrado.

La gravedad de los errores en la categoría 4 pueden ser muy variada. En un caso extremo, la traducción de una oración produce una estructura sintáctica muy compleja o un estilo discursivo diferente al estilo discursivo natural en español.

Los errores en las categorías 5, 6 y 7 pueden afectar seriamente la validez del reactivo, ya que involucran, respectivamente, la cantidad, la calidad y la precisión de la información que recibe el estudiante. En general, estos errores pueden tener un impacto en la validez del reactivo en su forma traducida.

Muestra de reactivos

Para los fines de este estudio se seleccionó al azar una muestra constituida exclusivamente por reactivos de opción múltiple, dos por cada área temática. La Tabla 1 muestra el número de áreas temáticas del examen y el número de reactivos que fueron evaluados..

Tabla I. Número de áreas temáticas y reactivos seleccionados por población y área de contenido

Pruebas	Áreas temáticas		Reactivos incluidos en la muestra	
	Población 1	Población 2	Población 1	Población 2
Ciencias Naturales	4	5	8	10
Matemáticas	6	6	12	12

Registro de Errores

Dos evaluadores registraron los errores observados en los reactivos en un formato de registro elaborado de acuerdo con la tipología de errores. Al identificar cada error se anotó la explicación del error o el sesgo indebido que sería de esperar debido a ese error, y que podría haber

favorecido o afectado negativamente a los alumnos evaluados con la versión en español de las pruebas. La Figura I muestra un reactivo en la versión en inglés y la versión en español. La Tabla II muestra el registro de errores de ese reactivo con el formato de registro empleado.

Figura 1. Reactivo J1, Población 2: versión en inglés y versión en español.

- J1. Which BEST describes the surface of the Earth over billions of years?
- A flat surface is gradually pushed up into higher and higher mountains until the Earth is covered with mountains.
 - High mountains gradually wear down until most of the Earth is at sea level.
 - High mountains gradually wear down as new mountains are continuously being formed, over and over again.
 - High mountains and flat plains stay side by side for billions of years with little change.
- J1. ¿Cuál de las siguientes afirmaciones describe MEJOR los cambios de la superficie de la Tierra a través de miles de millones de años?
- Una superficie plana se transforma gradualmente en montañas altas hasta que la Tierra es cubierta con montañas.
 - Las montañas muy altas fueron bajando gradualmente hasta que la mayor parte de la Tierra bajó al nivel del mar.
 - Las montañas muy altas fueron transformándose gradualmente en superficies planas que quizás se transformaron otra vez en montañas y así sucesivamente.
 - Las montañas muy altas y las superficies planas se mantuvieron por miles de millones de años con pequeños cambios.

Tabla II. Ejemplo de formato de registro de errores: reactivo J1, Población 2

Reactivo	Tipo de error	Descripción	Sesgo
J1	1. Estilo y ortografía		
	2. Formato	* Margen derecho justificado	
	3. Convenciones	* Inconsistencia de tiempos gramaticales entre las opciones	
	4. Gramática y sintaxis		
	5. Información	* La noción de que las montañas se desgastan desaparece en la traducción * La opción C agrega indebidamente "quizás" * Opción C: "high mountains" traducido como "las montañas más altas" * Interpretación indebida de "...as new mountains..." (opción C)	<
	6. Semántica	* Interpretación incorrecta de "wear down" como "were down" (opciones B y C)	<

Reactivo	Tipo de error	Descripción	Sesgo
		* Al usar indebidamente el tiempo pretérito, las opciones dejan de reflejar el sentido de los procesos de la Tierra, que son continuos * Omisión de "...side by side..." en la traducción; traducción incorrecta de "flat plain" (opción D)	
	7. Contenido	* La oración base agrega el término "cambios" (de la superficie de la Tierra)	

Nota: Un sesgo probable a favor o en contra de los estudiantes a quienes se aplicó la prueba en español se denotó, respectivamente, con los símbolos ">" y "<".

Resultados

Debido a la naturaleza exploratoria de este estudio, el análisis de los datos se limitó a examinar: 1) la frecuencia con que se observó cada tipo de error y 2) el porcentaje de reactivos con errores serios de gramática y sintaxis y con errores de información, semántica y contenido. Estos análisis se efectuaron por área de contenido y por población.

La Tabla III presenta el número de errores identificados en los reactivos seleccionados por cada área de contenido y cada población. Se observaron más errores en los reactivos de la Población 2 (43 errores) que en los reactivos de la Población 1 (27 errores). La diferencia no parece deberse a que el número de reactivos seleccionados sea mayor para la Población 2 (22 reactivos) que para la Población 1 (20 reactivos). Esta diferencia puede deberse, más bien, a que el texto de los reactivos de la Población 2 tiende a ser más largo y más complejo que los reactivos de la Población 1.

Tabla III. Frecuencia de tipos de errores de los reactivos por población y área de contenido

Categoría	Población 1			Población 2		
	Ciencias naturales (8 reactivos)	Matemáticas (12 reactivos)	Total (20 reactivos)	Ciencias naturales (10 reactivos)	Matemáticas (12 reactivos)	Total (22 reactivos)
1. Estilo y ortografía	1	0	1	1	1	2
2. Formato	0	1	1	3	1	4
3. Convenciones	3	3	6	9	3	12
4. Gramática y sintaxis	3	5	8	5	0	5
5. Información	0	4	4	6	4	10
6. Semántica	0	5	5	5	2	7
7. Contenido	1	1	2	3	0	2
Total	8	19	27	32	11	43
Porcentaje de errores	30	70	100	74	26	100

El análisis del porcentaje de errores por área de contenido revela distintas tendencias entre los reactivos de las dos poblaciones. Para la Población 1, el 30% de los errores identificados fue en reactivos de ciencias naturales y el 70% de los errores identificados fue en reactivos de matemáticas. Para la Población 2, esta proporción se invierte: 74% de los errores identificados

proviene de reactivos de ciencias naturales mientras que 26% proviene de reactivos de matemáticas.

Los errores observados son más frecuentes en las categorías 4 a 7, que son los errores más graves, aquéllos que generalmente pueden atentar contra la validez de la prueba. Para la Población 1, del total de 27 errores observados, 8, 4, 5 y 2 pertenecen respectivamente a las categorías Gramática y sintaxis, Información, Semántica y Contenido. Para la Población 2, de los 43 errores observados, 5, 10, 7 y 2 pertenecen respectivamente a esas categorías. En otras palabras, 70% de los errores observados entre los reactivos de la Población 1 y 56% de los errores observados entre los reactivos de la Población 2 pertenecen a categorías que tienden a atentar contra la validez de los reactivos.

El hecho de que puede haber varios errores en un mismo reactivo, dificulta estimar el impacto que ellos pueden tener en una prueba. Por ejemplo, si todos los errores están concentrados en unos cuantos reactivos, el impacto potencial de una traducción inadecuada podría considerarse como menor que si estos errores estuvieran distribuidos entre los reactivos. En consecuencia, a fin de apreciar con mayor precisión el impacto de los errores de traducción observados, se hizo un análisis adicional para identificar cuántos reactivos tienen errores serios que podrían atentar contra la validez de la prueba en español.

La Tabla IV presenta de manera sintética estos resultados. De la muestra de 42 reactivos seleccionados para este estudio, 15 reactivos (9 de la Población 1 y 6 de la Población 2) son graves. En otras palabras, el 36% de los reactivos presenta errores que podrían ser considerados como errores graves pueden producir un sesgo.

Tabla IV. Cantidad de reactivos con errores serios.

Población	Áreas temáticas	Reactivos		
		Seleccionados	Con error	% con error
1	Ciencias Naturales	8	2	25.0
	Matemáticas	12	7	58.3
2	Ciencias	10	3	33.3
	Matemáticas	12	3	25.0
Total		42	15	35.7

Es importante hacer notar que entre los reactivos identificados como con errores serios, se identificó uno que podría producir un sesgo en favor de los estudiantes evaluados con la prueba en español. La oración base usa el término “desigualdad” para referirse a un conjunto de desigualdades que el alumno tiene que analizar. Dicho término no aparece en la versión en inglés. El uso de dicho término en la versión en español puede proporcionar al alumno información clave para resolver el problema. Como se puede apreciar, los errores de traducción pueden producir tanto sesgos a favor como en contra de los estudiantes a quienes se evalúa con la traducción de las pruebas. Ambos sesgos, sin embargo, son igualmente inaceptables.

Conclusiones

Los resultados de este estudio exploratorio justifican la necesidad de analizar la totalidad de los reactivos TIMSS-1995-México con una versión perfeccionada de la tipología y con procedimientos más rigurosos en los que existan jueces independientes y se estime la confiabilidad entre observadores.

Un hecho que muestra la importancia de realizar estudios como éste tiene que ver con las mismas limitaciones de los métodos existentes para traducir pruebas y evaluar la calidad de las traducciones de pruebas. Uno de los informes sobre la calidad de los datos de TIMSS incluye un capítulo sobre los procedimientos de revisión empleados en TIMSS-1995 (Mullis, Kelly y Haley, 1996) según el cuál, después de aplicar el riguroso proceso de revisión, no se identificó ningún reactivo problemático en las versiones finales de los cuadernillos empleados en México.

Tal inconsistencia sugiere que muchos y muy serios errores de traducción de pruebas pueden pasar desapercibidos a pesar de los procedimientos sofisticados empleados por TIMSS en 1995 (o incluso aquéllos diseñados posteriormente para TIMSS 1999-R y PISA) si no se les implementa adecuadamente. El proceso de traducción de las pruebas TIMSS-1995 se basó en gran medida en recomendaciones que posiblemente tuvieron un gran margen de interpretación por parte de los representantes de los países participantes.

En el momento de escribir este informe, no se cuenta con datos sobre las características de las personas que tradujeron los reactivos de TIMSS-1995. Tampoco se cuenta con documentación sobre los procedimientos de traducción empleados en México. De tal suerte, se desconoce el grado de precisión con que se implementaron los lineamientos y recomendaciones de TIMSS. Por lo tanto, aunque el análisis de los factores lingüísticos de este estudio permite identificar deficiencias en la traducción de los reactivos TIMSS-1995, no es posible discernir si esas deficiencias se deben a una selección inadecuada de los traductores, a que no se empleó o no se implementaron debidamente los procedimientos recomendados por TIMSS, o a que el sistema de revisión de las traducciones de TIMSS tiene fallas que impidieron detectar y corregir esos errores. Sin embargo, basándose en los resultados de estudios sobre la calidad de la traducción de los reactivos y los procedimientos de traducción empleados para evaluar otras pruebas (Grisay, 1998), es posible que una combinación de los tres factores hayan contribuido a que se dieran esas deficiencias.

Este ha sido un estudio exploratorio. Las conclusiones aquí expresadas requieren el respaldo del estudio completo, en el que se analice la totalidad de los reactivos TIMSS-1995 y se emplee una tipología más elaborada y un procedimiento de evaluadores independientes más riguroso. Es razonable esperar que a partir de los resultados de ese estudio final será posible establecer con precisión la forma en que México implementará los procedimientos de traducción de instrumentos en futuras comparaciones internacionales.

Bibliografía

- American Translators Association: Framework for standard error marking and explanation. <http://www.atanet.org> retrieved October 10, 2003.
- American Educational Research Association. (1999). *Standards for Educational and Psychological Testing*. Washington, DC: Author.
- Baxter, G. P., Shavelson, R. J., Goldman, S. R., y Pine, J. (1992). Evaluation of procedure-based scoring for hands-on science assessment. *Journal of Educational Measurement*, 29(1), 1-17.
- Camilli, G. y Shepard, L. A. (1994). *Methods for identifying biased items*. Thousand Oaks, California: Sage.
- Ercikan, K. (1998). Translation effects in international assessment. *International Journal of Educational Research*, 29, 543-553.
- Grisay, A. (1998). Translation and cultural appropriateness of the test and survey material.
- Hambleton, R.K. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment*, 10(3), 229-244.
- Hambleton, R.K., & Jones, R.W. (1994). Comparisons of empirical and judgemental methods for detecting differential item functioning. *Educational Research Quarterly*, 18(1), 21-36.

- Maxwell, B. (1996). Translation and cultural adaptation of the survey instruments. En Martin, M.O. y Kelly, D.L. (eds). *Third International Mathematics and Science Study (TIMSS Technical Report, Volume 1: Design and Development*. Chestnut Hill, MA: Boston College.
- Messick, S. (1995). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice*, 14(4), 5–8.
- Mullis, I.V.S., Kelly, D.L., and Haley, K. (1996). "Translation Verification Procedures" in M.O. Martin and I.V.S. Mullis *Third International Mathematics and Science Study: Quality Assurance in Data Collection*. Chestnut Hill, MA: Boston College.
- Solano-Flores, G. (2003). The multidimensionality of test review and test design: A conceptual framework for addressing linguistic and cultural diversity in testing. Paper accepted at the 10th Biennial Conference of the European Association for Research on Learning and Instruction, Padova, Italy - August 26 – 30, 2003.
- Solano-Flores, G., & Shavelson, R. J. (1997). Development of performance assessments in science: Conceptual, practical, and logistical issues. *Educational Measurement: Issues and Practice*, 16(3), 16-25.
- Solano-Flores, G., & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English-language learners. *Educational Researcher*, 32(2), 3-13.
- Solano-Flores, G., Trumbull, E., & Nelson-Barber, S. (2002). Concurrent Development of Dual Language Assessments: An Alternative to Translating Tests for Linguistic Minorities. *International Journal of Testing*, 2(2), 107-129.
- Solano-Flores, G., & Nelson-Barber, S. (2001). On the cultural validity of science assessments. *Journal of Research in Science Teaching*, 38(5), 553-573.
- Valdés, G., & Figueroa, R. A. (1994). *Bilingualism and testing: A special case of bias*. Norwood, NJ: Ablex.
- Van de Vivjer, F., & Poortinga, Y. H. (1997). Towards an integrated analysis of bias in cross-cultural assessment. *European Journal of Psychological Assessment*, 13(1), 29-37.
- Van de Vivjer, F., & Tanzer, N. K. (1998). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*, 47(4), 263-279.

Apéndice I:

Tipología de Errores

1. Estilo y ortografía

La traducción del reactivo no es consistente con las convenciones estilísticas y ortográficas utilizadas en textos impresos en el país. El rango de la severidad de este tipo de errores es muy amplio, desde errores que pueden pasar desapercibidos por muchos lectores hasta errores que pueden llegar a afectar la dificultad de la lectura. Esta categoría incluye errores como:

- uso inadecuado de acentos y puntuación diacrítica
- uso inadecuado de mayúsculas y minúsculas
- inconsistencia gramatical entre el sujeto y el verbo de una oración

2. Formato

La información visual del reactivo en español se presenta de manera diferente al reactivo en el idioma original. Esta información visual incluye tablas, gráficas e ilustraciones. Algunos errores de formato pueden tener su origen en fallas en la forma de transmitir, copiar o procesar los componentes gráficos de un reactivo.

Otros errores de formato pueden deberse a diferencias entre idiomas en el espacio que ocupa el texto impreso que deben considerarse desde etapas iniciales en el diseño de pruebas internacionales. Por ejemplo, las palabras y oraciones en español tienden a ser más largas que en inglés. Para poder cumplir con el mismo número de reactivos por página que en el idioma original, el diseño de la prueba en español puede llegar a sacrificar algo del espacio de respuesta en reactivos de respuesta abierta o reducir el tamaño de los márgenes o de la tipografía. Tales cambios pueden incrementar la dificultad del reactivo por razones diferentes al dominio del conocimiento que la prueba evalúa.

Esta categoría incluye errores como:

- ubicación o tamaño diferente de tablas, gráficas o ilustraciones
- cambios en el tamaño o tipo de letra
- variación de la ilustración

3. Convenciones

La traducción no refleja el estilo convencional de reactivos en el idioma español o viola principios de redacción de reactivos aceptados universalmente. Errores de este tipo pueden incrementar o reducir la demanda cognoscitiva del reactivo. Esta categoría incluye errores como:

- inconsistencia gramatical entre la oración raíz y las opciones en un reactivo de opción múltiple
- uso inadecuado de puntos suspensivos para denotar continuidad entre la oración raíz y las opciones
- inconsistencia gramatical entre las opciones
- inconsistencia sintáctica entre la oración y las opciones

4. Gramática y sintaxis

La estructura gramatical del reactivo no es tan natural como si éste hubiera sido escrito originalmente en español o no es adecuada para la edad del estudiante examinado. Este tipo de error afecta la claridad del reactivo, cuya estructura discursiva puede ser diferente de la usual en el idioma español. Además, este tipo de error puede llegar a incrementar innecesariamente la demanda de lectura del reactivo. Esta categoría incluye errores como:

- preposiciones incorrectas
- uso inapropiado del plural en español
- sintaxis más compleja que en el original
- sintaxis incorrecta
- apego literal a la forma sintáctica en la lengua original (traducción palabra por palabra)

5. Información

La traducción altera la cantidad o el contenido de la información que se proporciona al estudiante en las instrucciones o el problema que se le presenta. Errores de este tipo pueden hacer que la

Población 1: Ciencia 1/2 dificultad de un reactivo sea menor o mayor en español que en el idioma original. Esta categoría incluye errores como:

- inserción u omisión de palabras, oraciones o explicaciones que no aparecen en el texto original
- omisión de palabras, oraciones o explicaciones
- uso de términos que no son los más apropiados para el nivel técnico del contenido del texto original
- inconsistencia en la manera en que se traduce un término en el texto
- omisión de un término importante para comprender el problema
- uso de una palabra difícil o poco común en español
- representación numérica no fiel a la versión en el idioma original

6. Semántica

Las ideas expresadas en el reactivo no tienen el mismo significado que en la lengua original debido a que la traducción es imprecisa. Errores graves de este tipo pueden conducir al alumno a cometer errores en la interpretación del reactivo. Esta categoría incluye errores como:

- uso de falsos cognatos
- uso innecesario de palabras difíciles o poco frecuentes
- uso de preposiciones imprecisas o inadecuadas o de tiempos verbales.
- interpretación incorrecta de palabras o de la sintaxis de las oraciones en el idioma original
- traducción imprecisa que altera la forma en que pueden interpretarse las opciones
- alteración del sentido original del texto

7. Contenido

La traducción altera el contenido del reactivo al grado de que puede llegar a afectar la equivalencia entre idiomas del constructo que se pretende medir. Errores de este tipo se deben a que el nivel del léxico o la formalidad del texto original se pierde en la traducción, lo que afecta al contenido mismo del reactivo. Esta categoría incluye errores como:

- falta de correspondencia de conceptos entre idiomas
- introducción de un término técnico que no aparece en el idioma original
- omisión de un término técnico
- inserción de término académico
- uso de un término académico incorrecto

Apéndice II: Registro de errores en la muestra de reactivos seleccionada

Población 1: Ciencia 1/2

N1			
	1. Estilo y ortografía		
	2. Fórmalo		
	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información		
N8	6. Semántica		
	7. Contenido		
	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones	* Uso indebido de puntos suspensivos en la oración base	
	4. Gramática y sintaxis		
04	5. Información		
	6. Semántica		
	7. Contenido		
	1. Estilo y ortografía	* "luna" no se escribe con mayúscula	
	2. Formato		
	3. Convenciones		
Q2	4. Gramática y sintaxis	* Uso incorrecto de "...para esto?"	
	5. Información		
	6. Semántica		
	7. Contenido		
	1. Estilo y ortografía		
	2. Formato		
Q7	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información		
	7. Contenido	* El término "experimento" no aparece en el original. * La situación descrita en el reactivo no se refiere en realidad a un experimento, sino una observación	<
	1. Estilo y ortografía		
	3. Convenciones	* Inconsistencia de tiempos gramaticales entre la oración base y las opciones * Uso indebido de puntos al final de las opciones	
Q9	4. Gramática y sintaxis	* Para expresar la idea de que se liberan burbujas de aire (Opción C), "las burbujas" debiera ser sujeto pasivo en vez de sujeto activo	<
	5. Información		
	6. Semántica		
	7. Contenido		

Población 1: Ciencia 2/2

R2	1. Estilo y ortografía		
	2. Formato		
	4. Gramática y sintaxis	* Anglicismo "ayudaría si tuviéramos..." (Opción C) en vez de "sería bueno tener..."	
	5. Información		
	6. Semántica		
	7. Contenido		
	R6	1. Estilo y ortografía	
2. Formato			
3. Convenciones		* Uso indebido de puntos al final de las opciones	
4. Gramática y sintaxis			
5. Información			
6. Semántica			

Población 1: Matemáticas ½

17	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información		
	6. Semántica		
	7. Contenido		
18	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información	* Los términos "área" y sombreado no aparecen en la versión en inglés	<
	6. Semántica	* La idea de que las figuras representan una fracción no está clara en la primera oración	<
	7. Contenido		
J1	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones	* Uso indebido de puntos en la oración base * Uso indebido de puntos al final de las opciones	
	4. Gramática y sintaxis		
	5. Información	* La primera oración no incluye el término "hexágono"	<
	6. Semántica		
	7. Contenido		
J4	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones		
	4. Gramática y sintaxis	* La complejidad sintáctica de la oración base dificulta la comparación de las operaciones * Sintaxis incorrecta en el planteamiento de una pregunta	<
	5. Información		
	6. Semántica		
	7. Contenido		
K4	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones		
	4. Gramática y sintaxis	Debiera decir "...la puntuación más alta..."	
	5. Información		
	6. Semántica		
	7. Contenido		
K7	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones		
	4. Gramática y sintaxis	* Estructura sintáctica compleja * Sintaxis incorrecta en el planteamiento de una pregunta	
	5. Información	* Omisión de "... (alambre) delgado"	
	6. Semántica	* La palabra "contomo" puede ser difícil	<
	7. Contenido		

Población 1: Matemáticas 2/2

K8	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones		
	4. Gramática v sintaxis		
	5. Información		
	6. Semántica		
	7. Contenido		
L4	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información		
	6. Semántica	* "pattern" traducido indebidamente como "orden"	<
	7. Contenido		
MI	1. Estilo y ortografía		
	2. Formato	;1; El uso de mayúsculas para hacer énfasis en ciertas palabras no es consistente con la versión en inglés	
	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información		
	6. Semántica	* "drops" traducido indebidamente como "lanza"	
	7. Contenido	* Se tradujo "best chance: como "más probable", que no aparece en inglés	<
M5	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información	* Se ha agregado una oración al principio del reactivo	
	6. Semántica		
	7. Contenido		
M6	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones	* Uso indebido de puntos al final de las opciones	
	4. Gramática y sintaxis	* Expresión incorrecta: "...¿Qué podría hacer al resultado...?"	
	5. Información		
	6. Semántica		
	7. Contenido		
M7	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información		
	6. Semántica	"1; La traducción incorrecta de "...wouid most likely..." como "mejor puede medirse en milímetros" produce que en español haya dos opciones correctas (A y C)	<

Población 2: Ciencia 1/2

I13	1. Estilo y ortografía		1
	2. Formato	"": Margen derecho justificado	
	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información		
	7. Contenido	* Uso del término "Celsius" en vez de "centígrado" * Omisión del término "precisión"	<
	J1	1. Estilo y ortografía	
2. Formato		* Margen derecho justificado	
3. Convenciones		sl; Inconsistencia de tiempos gramaticales entre las opciones	
5. Información		* La noción de que las montañas se desgastan desaparece en la traducción * La opción C agrega indebidamente "quizás" * Opción C: "high mountains" traducido como "las montañas más altas" * Interpretación indebida de "...as new mountains..." (opción C)	
6. Semántica		* Interpretación incorrecta de "wear down" como "were down" (opciones B y C) * Al usar indebidamente el tiempo pretérito, las opciones dejan de reflejar el sentido de los procesos de la Tierra, que son continuos * Omisión de "...side by side..." en la traducción; traducción incorrecta de "fiat plain" (opción D)	<
7. Contenido		* La oración base agrega el término "cambios (de la superficie de la Tierra)"	<
J8		1. Estilo y ortografía	
	2. Formato		
	3. Convenciones	* Uso indebido de puntos al final de las opciones	
	4. Gramática y sintaxis		
	5. Información		
	6. Semántica		
	7. Contenido		
K11	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información		
	6. Semántica		
	7. Contenido		
K17	1. Estilo y ortografía	* Coma innecesaria en "de las tres posiciones,..." en la oración base	
	2. Formato		
	3. Convenciones	* Insistencia de forma gramatical ("en la") entre las opciones * Uso indebido de puntos al final de las opciones	
	4. Gramática y sintaxis	* Estructura inusual ("..la gravedad actúa" en vez de "actúa la gravedad") en forma interrogativa * Opciones B y C: la palabra "posición" debiera estar en plural	
	5. Información	* La pregunta en la oración base debiera ser "¿En cuáles posiciones..." (no "¿En cuál de las tres posiciones...") para asegurar la plausibilidad de todas las opciones	<
	6. Semántica		
	7. Contenido		
N4	1. Estilo y ortografía		
	2. Formato	* Margen derecho justificado	
	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información		
	6. Semántica		
	7. Contenido		

Población 2: Ciencia 2/2

N5	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones	* Inconsistencia de tiempos gramaticales entre la oración base y las opciones * Inconsistencia gramatical en número entre la oración base y la opción A * Uso indebido de puntos al final de las opciones	
	4. Gramática y sintaxis	* Voz pasiva de "bombeados" poco natural en español	
	5. Información	si: Se agregó "La unión (de gases)"	
	6. Semántica	* Traducción literal de "pnmped" como "bombeados"	
	7. Contenido		
Q15	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones	* Uso indebido de puntos al final de las operaciones	
	4. Gramática y sintaxis	* Uso del gerundio resultado de una traducción literal de las opciones * "¿.Cuál de los siguientes..." es gramaticalmente incorrecto	
	5. Información		
	6. Semántica	* Traducción literal de "baking bread" como "horneando pan" en vez de "pan homeándose"	
	7. Contenido		
Q16	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones	* Uso indebido de puntos al final de las opciones	
	4. Gramática y sintaxis		
	5. Información		
	6. Semántica		
	7. Contenido		
RI	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información		
	6. Semántica		
	7. Contenido		

Población 2: Matemáticas ½

K4	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones		
	4. Gramática v sintaxis		
	5. Información	* En la primera oración se agrega el término "desigualdad"	
	6. Semántica		
	7. Contenido		
L10	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones	* Uso indebido de puntos al final de las opciones	
	4. Gramática y sintaxis		
	5. Información		
	6. Semántica		
	7. Contenido		
L12	1. Estilo y ortografía	* Uso de "Cada quien" en vez de "Cada uno"	
	2. Formato		
	3. Convenciones	* Uso indebido de puntos al final de las opciones	
	4. Gramática y sintaxis		
	5. Información		
	6. Semántica		
	7. Contenido		
L14	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información		
	6. Semántica		
	7. Contenido		
L15	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones	* Uso indebido de puntos al final de las opciones	
	4. Gramática y sintaxis		
	5. Información		
	6. Semántica		
	7. Contenido		
02	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información		
	6. Semántica		
	7. Contenido		

Población 2: Matemáticas 2/2

08	1. Estilo y ortografía		
	2. Fórmalo		
	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información	* La primera oración no proporciona un término alternativo para la noción de rotar	<
	6. Semántica	* La traducción usa "centro de giro" en vez de "centro de rotación"	
	7. Contenido		<
P11	1. Estilo y ortografía		
	2. Fórmalo	* Ligera variación en la ilustración (se agregó un borrador al lápiz)	
	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información		
	6. Semántica		
	7. Contenido		
P17	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información		
	6. Semántica		
	7. Contenido		
R7	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información		
	6. Semántica	* Uso innecesario de "grosor" en vez de "grueso"	
	7. Contenido		
R9	1. Estilo y ortografía		
	2. Formato		
	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información	* La oración base agrega el término "igualdades"	
	6. Semántica		
	7. Contenido		
V3	1. Estilo y ortografía		1
	2. Formato		
	3. Convenciones		
	4. Gramática y sintaxis		
	5. Información	* Se eliminó la palabra "total"	
	6. Semántica		
	7. Contenido		<

Apéndice III:

Reactivos incluidos en la muestra en inglés y en español

K4. $\frac{x}{2} < 7$ is equivalent to

A. $x < \frac{7}{2}$

B. $x < 5$

C. $x < 14$

D. $x > 5$

E. $x > 14$

N4. Hace años los agricultores encontraron que las plantas de maíz crecían mejor si se enterraba pescado en mal estado cerca de éstas. ¿Qué es lo que probablemente aporta el pescado en mal estado que favorece el crecimiento de las plantas?

- A. Energía
- B. Minerales
- C. Proteínas
- D. Oxígeno
- E. Agua

N5. One of the principal causes of acid rain is

- A. waste acid from chemical factories being pumped into rivers
- B. acid from chemical laboratories evaporating into the air
- C. gases from burning coal and oil dissolving in water in the atmosphere
- D. gases from air conditioners and refrigerators escaping into the atmosphere

N5. Una de las principales causas de la lluvia ácida es:

- A. Los desechos ácidos de fábricas de productos químicos que son bombeados a los ríos.
- B. El ácido de laboratorios químicos que se evapora en el aire.
- C. La unión de los gases de la combustión del carbón y el petróleo con el agua que existe en la atmósfera.
- D. El gas de aparatos de aire acondicionado y refrigeradores que escapa hacia la atmósfera.

Q15. Which is NOT an example of a chemical change?

- A. Boiling water
- B. Rusting iron
- C. Burning wood
- D. Baking bread

Q-15

Q15. ¿Cuál de los siguientes NO es un ejemplo de un cambio químico?

- A. Agua hirviendo.
- B. Fierro oxidándose.
- C. Madera quemándose.
- D. Horneando pan.

Q16. How long does it take light from the nearest star other than the Sun to reach Earth?

- A. Less than 1 second
- B. About 1 hour
- C. About 1 month
- D. About 4 years

Q-16

Q16. ¿Cuánto tiempo le toma a la luz de la estrella más cercana (diferente del Sol) llegar a la Tierra?

- A. Menos de 1 segundo.
- B. Cerca de 1 hora.
- C. Cerca de 1 mes.
- D. Cerca de 4 años.

K4. La desigualdad $\frac{x}{2} < 7$ es equivalente a...

A. $x < \frac{7}{2}$

B. $x < 5$

C. $x < 14$

D. $x > 5$

E. $x > 14$

L10. This chart shows temperature readings made at different times on four days.

TEMPERATURES					
	6 a.m.	9 a.m.	Noon	3 p.m.	8 p.m.
Monday	15°	17°	20°	21°	19°
Tuesday	15°	15°	15°	10°	9°
Wednesday	8°	10°	14°	13°	15°
Thursday	8°	11°	14°	17°	20°

When was the highest temperature recorded?

- A. Noon on Monday
- B. 3 p.m. on Monday
- C. Noon on Tuesday
- D. 3 p.m. on Wednesday

10. Este cuadro muestra algunas lecturas de temperaturas tomadas en cuatro días a diferente hora.

TEMPERATURAS					
	6 a.m.	9 a.m.	12 mediodía	3 p.m.	8 p.m.
Lunes	15°	17°	20°	21°	19°
Martes	15°	15°	15°	10°	9°
Miércoles	8°	10°	14°	13°	15°
Jueves	8°	11°	14°	17°	20°

¿Cuándo se registró la temperatura más alta?

- A. Al mediodía del lunes.
- B. A las 3 p.m. del lunes.
- C. Al mediodía del martes.
- D. A las 3 p.m. del miércoles.

L12. Four children measured the width of a room by counting how many paces it took them to cross it. The chart shows their measurements.

Who had the longest pace?

- A. Stephen
- B. Erlane
- C. Ana
- D. Carlos

Name	Number of Paces
Stephen	10
Erlane	8
Ana	9
Carlos	7

12. Cuatro niños midieron el ancho de una habitación contando cuántos pasos dieron para atravesarla. El cuadro siguiente muestra los pasos que dio cada quien.

¿Quién dio los pasos más largos?

- A. Esteban
- B. Eduardo
- C. Ana
- D. Carlos

Nombre	Número de pasos
Esteban	10
Eduardo	8
Ana	9
Carlos	7

L14. The table shows the values of x and y , where x is proportional to y .

x	3	6	P
y	7	Q	35

What are the values of P and Q ?

- A. $P = 14$ and $Q = 31$
- B. $P = 10$ and $Q = 14$
- C. $P = 10$ and $Q = 31$
- D. $P = 14$ and $Q = 15$
- E. $P = 15$ and $Q = 14$

L14. La tabla muestra los valores de X y de Y , donde X es proporcional a Y .

X	3	6	P
Y	7	Q	35

¿Cuáles son los valores de P y Q ?

- A. $P = 14$ y $Q = 31$
- B. $P = 10$ y $Q = 14$
- C. $P = 10$ y $Q = 31$
- D. $P = 14$ y $Q = 15$
- E. $P = 15$ y $Q = 14$

L15. In a quadrilateral, two of the angles each have a measure of 110° , and the measure of a third angle is 90° . What is the measure of the remaining angle?

- A. 50°
- B. 90°
- C. 130°
- D. 140°
- E. None of the above

15. En un cuadrilátero, dos de los ángulos miden 110° cada uno, y la medida del tercero es de 90° . ¿Cuál es la medida del otro ángulo?

- A. 50°
- B. 90°
- C. 130°
- D. 140°
- E. Ninguna de las anteriores.

O2. If the price of a can of beans is raised from 60 cents to 75 cents, what is the percent increase in the price?

A. 15%

B. 20%

C. 25%

D. 30%

12. Si el precio de una lata de frijoles ha subido de 60 centavos a 75 centavos, ¿en qué porcentaje ha aumentado el precio?

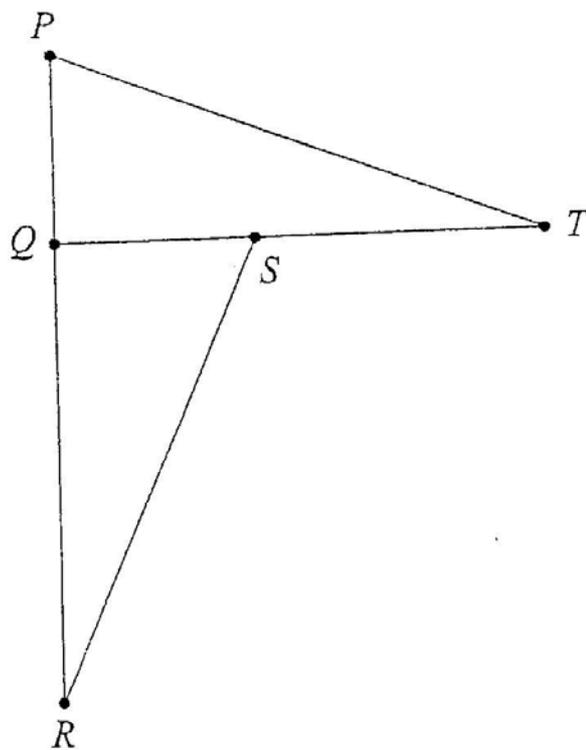
A. 15%

B. 20%

C. 25%

D. 30%

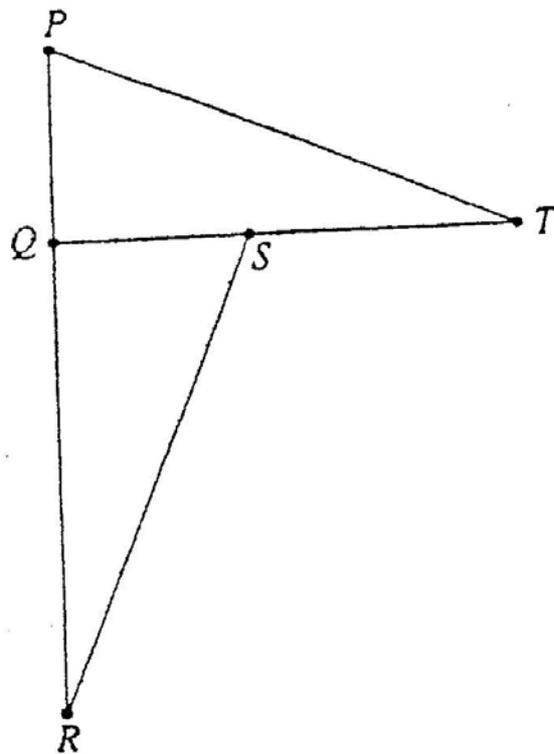
O8. Triangle PQT can be rotated (turned) onto triangle SQR .



What point is the center of rotation?

- A. P
- B. Q
- C. R
- D. S
- E. T

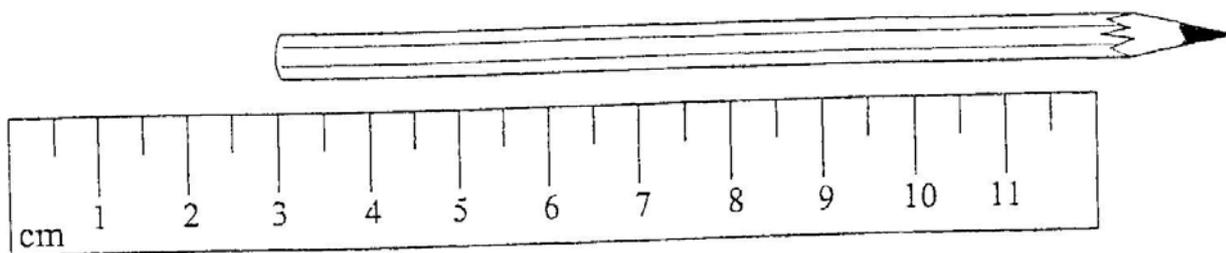
08. El triángulo PQT se puede girar para convertirse en el triángulo SQR .



¿Qué punto es el centro del giro?

- A. P
- B. Q
- C. R
- D. S
- E. T

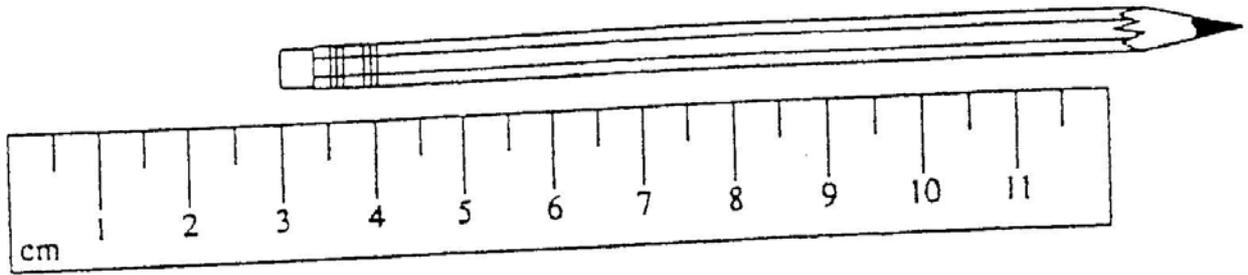
P11.



Which of these is closest to the length of the pencil in the figure?

- A. 9 cm
- B. 10.5 cm
- C. 12 cm
- D. 13.5 cm

P11.



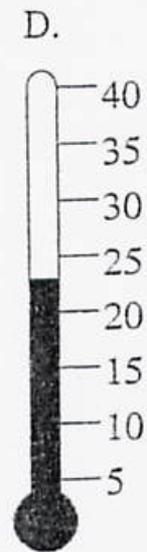
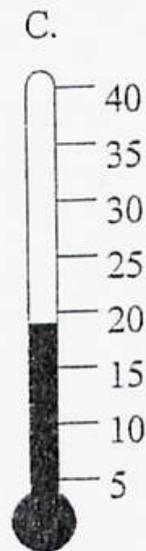
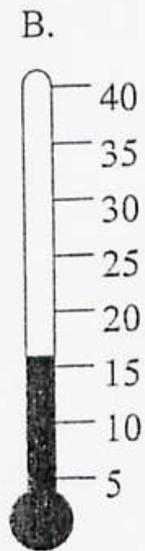
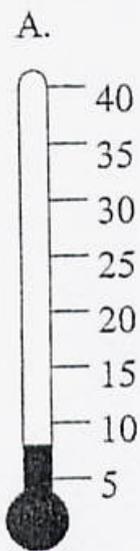
¿Cuál es la medida más aproximada de este lápiz?

- A. 9 cm
- B. 10.5 cm
- C. 12 cm
- D. 13.5 cm

P17. This table shows temperatures at various times during the week.

TEMPERATURES					
	6 a.m.	9 a.m.	Noon	3 p.m.	8 p.m.
Monday	15°	17°	20°	21°	19°
Tuesday	15°	15°	15°	10°	9°
Wednesday	8°	10°	14°	13°	15°
Thursday	8°	11°	14°	17°	20°

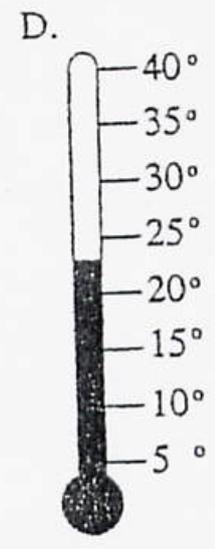
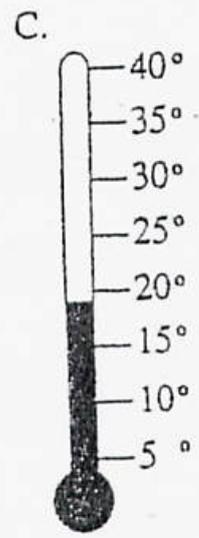
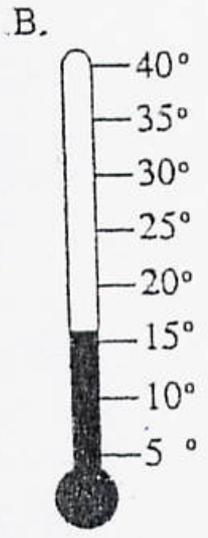
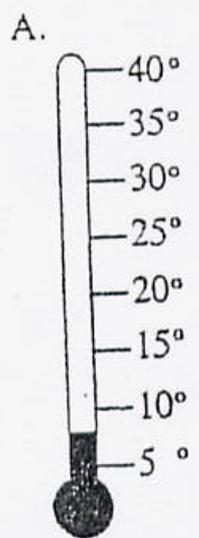
Which thermometer shows the temperature at 8 p.m. on Monday?



17. Este cuadro muestra algunas lecturas de temperaturas tomadas en cuatro días a diferente hora.

TEMPERATURAS					
	6 a.m.	9 a.m.	medio día	3 p.m.	8 p.m.
Lunes	15°	17°	20°	21°	19°
Martes	15°	15°	15°	10°	9°
Miércoles	8°	10°	14°	13°	15°
Jueves	8°	11°	14°	17°	20°

¿Qué termómetro muestra la temperatura del lunes a las 8 p.m.?



R7. A stack of 200 identical sheets of paper is 2.5 cm thick. What is the thickness of one sheet of paper?

- A. 0.008 cm
- B. 0.0125 cm
- C. 0.05 cm
- D. 0.08 cm

R7. Un paquete de 200 hojas de papel idénticas tiene 2.5 cm de grosor. ¿Cuál es el grosor de una hoja de papel?

A. 0.008 cm

B. 0.0125 cm

C. 0.05 cm

D. 0.08 cm

R9. Which one of the following is FALSE when a , b , and c are different real numbers?

A. $(a + b) + c = a + (b + c)$

B. $ab = ba$

C. $a + b = b + a$

D. $(ab)c = a(bc)$

E. $a - b = b - a$

R9. ¿Cuál de las siguientes igualdades es FALSA si a , b y c son números reales distintos?

A. $(a + b) + c = a + (b + c)$

B. $ab = ba$

C. $a + b = b + a$

D. $(ab)c = a(bc)$

E. $a - b = b - a$

V3. To mix a certain color of paint, Alana combines 5 liters of red paint, 2 liters of blue paint, and 2 liters of yellow paint. What is the ratio of red paint to the total amount of paint?

A. $\frac{5}{2}$

B. $\frac{9}{4}$

C. $\frac{5}{4}$

D. $\frac{5}{9}$

V3. Para obtener una mezcla de un determinado color, Ana combina 5 litros de pintura roja, 2 litros de pintura azul y 2 litros de pintura amarilla. ¿Cuál es la proporción de pintura roja en la mezcla?

A. $\frac{5}{2}$

B. $\frac{9}{4}$

C. $\frac{5}{4}$

D. $\frac{5}{9}$

N1. This table shows the temperature and precipitation (rain or snow) in four different towns on the same day.

	Town A	Town B	Town C	Town D
Lowest Temperature	13°C	-9°C	22°C	-12°C
Highest Temperature	25°C	-1°C	30°C	-4°C
Precipitation (rain or snow)	0 cm	5 cm	2.5 cm	0 cm

Where did it snow?

- A. Town A
- B. Town B
- C. Town C
- D. Town D

11. El siguiente cuadro muestra información acerca de la temperatura y la precipitación (lluvia o nieve) en cuatro diferentes ciudades el mismo día.

	Ciudad A	Ciudad B	Ciudad C	Ciudad D
Temperatura más baja	13°C	-9°C	22°C	-12°C
Temperatura más alta	25°C	-1°C	30°C	-4°C
Precipitación (lluvia o nieve)	0 cm	5 cm	2.5 cm	0 cm

¿Dónde nevó?

- A. Ciudad A
- B. Ciudad B
- C. Ciudad C
- D. Ciudad D

N8. A powder made up of both white specks and black specks is likely to be

- A. a solution
- B. a pure compound
- C. a mixture
- D. an element

N8. Un polvo hecho de partículas blancas y negras es probable que sea...

- A. una solución.
- B. un compuesto puro.
- C. una mezcla.
- D. un elemento.

O4. The Moon produces no light, and yet it shines at night. Why is this?

- A. The Moon reflects the light from the Sun.
- B. The Moon rotates at a very high speed.
- C. The Moon is covered with a thin layer of ice.
- D. The Moon has many craters.

La Luna no produce luz, y aun así brilla de noche. ¿Cuál es la mejor explicación para esto?

- A. La Luna refleja la luz del Sol.
- B. La Luna gira a una velocidad muy alta.
- C. La Luna está cubierta con una capa delgada de hielo.
- D. La Luna tiene muchos cráteres.

Q2. What is the BEST reason for including fruits and leafy vegetables in a healthy diet?

- A. They have a high water content.
- B. They are the best source of protein.
- C. They are rich in minerals and vitamins.
- D. They are the best source of carbohydrates.

Q2. ¿Cuál es la MEJOR razón para incluir frutas y vegetales verdes en una dieta saludable?

- A. Tienen un alto contenido de agua.
- B. Son la mejor fuente de proteínas.
- C. Son ricos en minerales y vitaminas.
- D. Son la mejor fuente de carbohidratos.

Q7. Four children can feel and smell an object inside a bag, but they cannot see it. Which of the following is NOT an observation about the object?

- A. "It is flat at one end and round at the other."
- B. "It smells like peppermint."
- C. "It has a bump on it."
- D. "I hope it is candy."

Q7. Cuatro niños pueden sentir y oler un objeto dentro de una bolsa, pero no pueden verlo. ¿Cuál de estas frases NO corresponde al experimento?

A. "Es plano en un extremo y redondo en el otro"

B. "Huele como a menta"

C. "Tiene un borde"

D. "Espero que sea un dulce"

Q9. Ken put a thermometer in a glass filled with hot water. Why does the liquid inside the thermometer rise?

- A. Gravity pushes it up.
- B. Air bubbles are released.
- C. Heat from the water makes it expand.
- D. Air pressure above the water pulls it up.

Q9. César colocó un termómetro dentro de un vaso lleno de agua caliente. ¿Por qué el líquido que está dentro del termómetro subió?

- A. La gravedad lo empuja hacia arriba.
- B. Las burbujas de aire se sueltan.
- C. El calor del agua lo hace subir.
- D. La presión del aire sobre el agua lo jala hacia arriba.

R2. Some children were trying to find out which of three light bulbs was brightest. Which one of these gives the best START toward finding the answer?

- A. "One bulb looks brightest to me, so I already know the answer."
- B. "All the bulbs look bright to me, so there cannot be an answer."
- C. "It would help if we had a way to measure the brightness of a light bulb."
- D. "We can take a vote and each person will vote for the bulb he or she thinks is the brightest."

2. Algunos niños estaban tratando de saber cuál de tres focos era el más brillante. ¿Cuál de estos enunciados es el mejor INICIO para encontrar la respuesta?
- A. "Un foco me parece más brillante, así que ya sé la respuesta"
 - B. "Todos los focos me parecen brillantes, así que no puede haber respuesta"
 - C. "Ayudaría si tuviéramos una forma de medir la brillantez de un foco"
 - D. "Podemos votar y cada persona votará por el foco que considere el más brillante"

R6. Which one of these groups contains only animals with a backbone (vertebrates)?

- A. Earthworm, snake, crayfish
- B. Spider, clam, dragonfly
- C. Crayfish, snake, horse
- D. Clam, fish, earthworm
- E. Horse, snake, fish

R6. ¿Cuál de estos grupos contiene sólo animales con columna vertebral (vertebrados)?

A. Lombriz, víbora, cangrejo.

B. Araña, almeja, libélula.

C. Cangrejo, víbora, caballo.

D. Almeja, pez, lombriz.

E. Caballo, víbora, pez.

17. Tanya has read the first 78 pages in a book that is 130 pages long. Which number sentence could Tanya use to find the number of pages she must read to finish the book?

A. $130 + 78 = \square$

B. $\square - 78 = 130$

C. $130 \div 78 = \square$

D. $130 - 78 = \square$

17. Antonio leyó las primeras 78 páginas de un libro de 130 páginas. ¿Qué operación aritmética debe usar Antonio para saber el número de páginas que le faltan de leer?

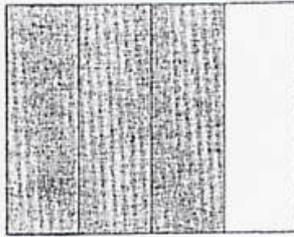
A. $130 + 78 = \square$

B. $\square - 78 = 130$

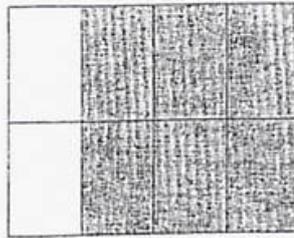
C. $130 \div 78 = \square$

D. $130 - 78 = \square$

18. Each figure represents a fraction.



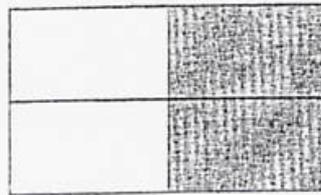
1



2



3

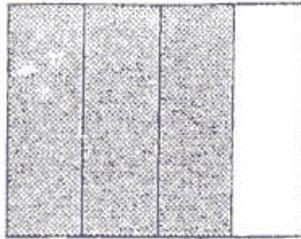


4

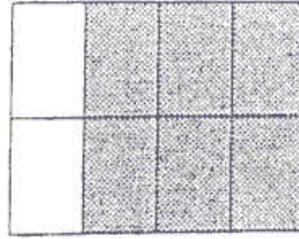
Which two figures represent the same fraction?

- A. 1 and 2
- B. 1 and 4
- C. 2 and 3
- D. 3 and 4

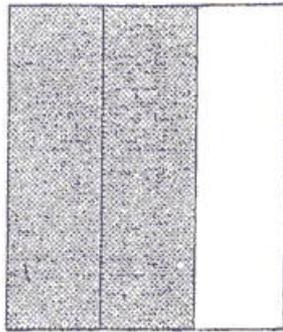
Cada figura tiene sombreada una fracción de su área.



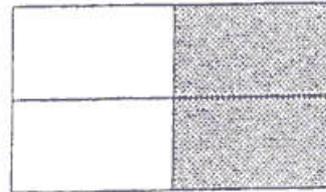
1



2



3

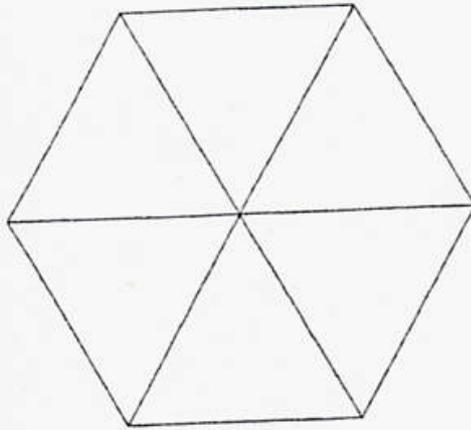


4

¿Cuáles son las dos figuras que tienen sombreada la misma fracción de su área?

- A. 1 y 2
- B. 1 y 4
- C. 2 y 3
- D. 3 y 4

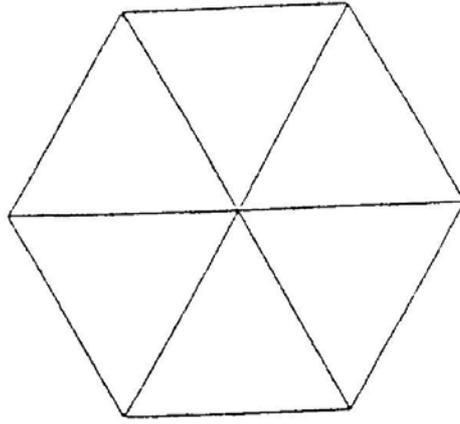
J1. Here is a hexagon.



The hexagon is divided into six

- A. triangles
- B. squares
- C. pentagons
- D. rectangles

Observa la siguiente figura.



El hexágono de la figura está dividido en seis...

- A. triángulos.
- B. cuadrados.
- C. pentágonos.
- D. rectángulos.

J4. 25×18 is more than 24×18 . How much more?

- A. 1
- B. 18
- C. 24
- D. 25

J4. El resultado de 25×18 , ¿cuánto es mayor que el resultado de 24×18 ?

A. 1

B. 18

C. 24

D. 25

K4. Kyle and Bob are playing a game. The object of the game is to get the highest total of points. This chart shows how many points they each scored.

Scorecard

Player	Kyle	Bob
Round 1	125	100
Round 2	125	125
Round 3	150	100
Round 4	50	150

Who won, and by how many points?

- A. Bob won by 25 points.
- B. Bob won by 100 points.
- C. Kyle won by 25 points.
- D. Kyle won by 175 points.

K4. José y Roberto están jugando. El objeto del juego es alcanzar la más alta puntuación en total. La tabla muestra los puntos que anotó cada jugador.

Puntuaciones

Jugador	José	Roberto
Partida 1	125	100
Partida 2	125	125
Partida 3	150	100
Partida 4	50	150

¿Quién ganó y por cuántos puntos?

- A. Roberto ganó por 25 puntos.
- B. Roberto ganó por 100 puntos.
- C. José ganó por 25 puntos.
- D. José ganó por 175 puntos.

K7. A thin wire 20 centimeters long is formed into a rectangle. If the width of this rectangle is 4 centimeters, what is its length?

- A. 5 centimeters
- B. 6 centimeters
- C. 12 centimeters
- D. 16 centimeters

K7. Si ocupamos 20 centímetros de alambre para hacer el contorno de un rectángulo que mide 4 centímetros de ancho, ¿cuánto medirá el largo de este rectángulo?

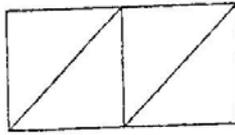
- A. 5 centímetros
- B. 6 centímetros
- C. 12 centímetros
- D. 16 centímetros

K8. Which rectangle is NOT divided into 4 equal parts?

A.



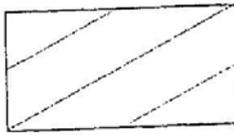
B.



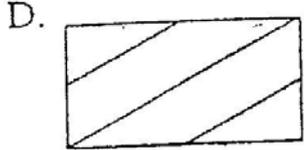
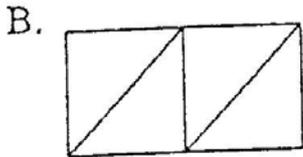
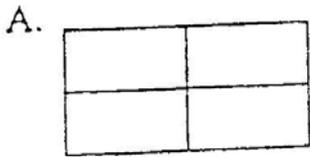
C.



D.



K8. ¿Cuál de estos rectángulos NO está dividido en cuatro partes iguales?



L4. These shapes are arranged in a pattern.



Which set of shapes is arranged in the same pattern?

- A. ★□★□★★□□★★□□
- B. □★□□★□□□★□□□□
- C. ★□★★□□★★★□□□
- D. □□★★□★□□★★□★

L4. Estas figuras están colocadas en orden.

○△○○△△○○○△△△

¿Cuál de los siguientes conjuntos de figuras tiene el mismo orden?

A. ★□★□★★□□★★□□

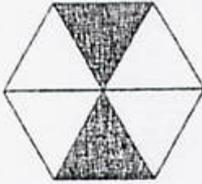
B. □★□□★□□□★□□□□

C. ★□★★□□★★★□□□

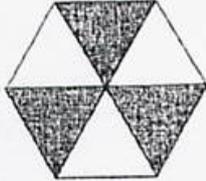
D. □□★★□★□□★★□★

M1. Samantha drops a stone onto each of these targets. The stone has the best chance of landing on a shaded space in which target?

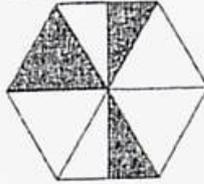
A.



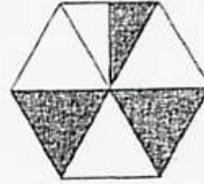
B.



C.

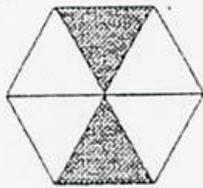


D.

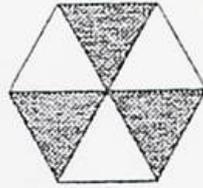


M1. Sara lanza una piedra a cada una de las figuras. ¿En cuál figura es MAS PROBABLE que la piedra caiga en un espacio sombreado?

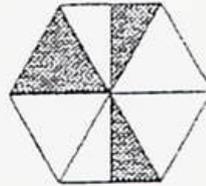
A.



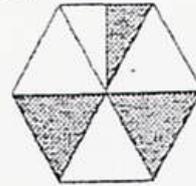
B.



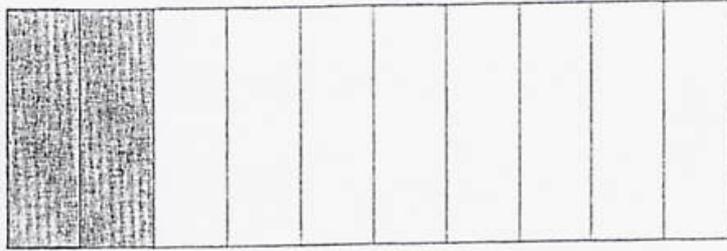
C.



D.



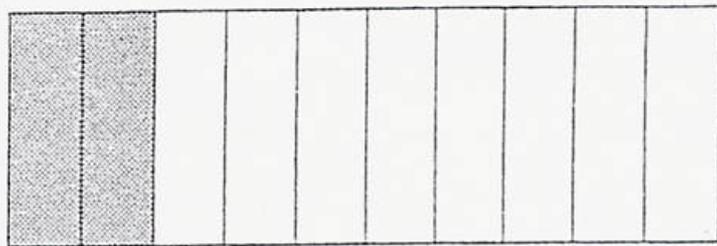
M5.



Which number represents the shaded part of the figure?

- A. 2.8
- B. 0.5
- C. 0.2
- D. 0.02

M5. Observa la siguiente figura.



¿Qué número representa la parte sombreada de la figura?

- A. 2.8
- B. 0.5
- C. 0.2
- D. 0.02

M6. John wanted to use his calculator to add 1463 and 319. He entered $1263 + 319$ by mistake. What could he do to correct his mistake?

- A. Add 200.
- B. Add 2.
- C. Subtract 2.
- D. Subtract 200.

M6. Juan estaba utilizando su calculadora para sumar 1463 y 319. Por error marcó $1263 + 319$. ¿Qué podría hacer al resultado para corregir su error?

- A. Sumar 200
- B. Sumar 2
- C. Restar 2
- D. Restar 200

M7. Which of these would most likely be measured in milliliters?

- A. The amount of liquid in a teaspoon
- B. The weight (mass) of a pin
- C. The amount of gasoline in a tank
- D. The thickness of 10 sheets of paper

- M7. ¿Cuál de los siguientes ejemplos es el que mejor puede medirse en mililitros?
- A. La cantidad de líquido en una cuchara pequeña.
 - B. El peso de un alfiler.
 - C. La cantidad de gasolina en un tanque.
 - D. El grueso de 10 hojas de papel.

113. The diagram shows five different Celsius thermometers. The body temperature of sick people ranges from about 36°C to 42°C . Which thermometer would be most suited for accurately measuring body temperature?

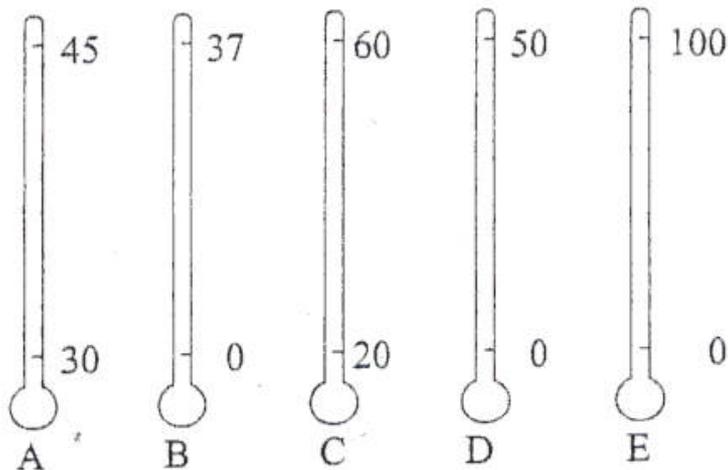
A. thermometer A

B. thermometer B

C. thermometer C

D. thermometer D

E. thermometer E



113. Los diagramas muestran cinco termómetros Celsius diferentes. La temperatura de una persona enferma varía entre los 36°C y los 42°C . ¿Qué termómetro es el más apropiado para medir la temperatura del cuerpo?

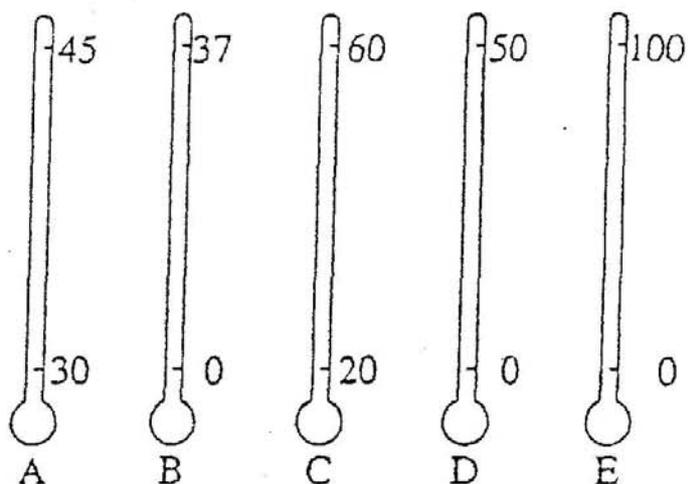
A. Termómetro A

B. Termómetro B

C. Termómetro C

D. Termómetro D

E. Termómetro E



- J1. Which BEST describes the surface of the Earth over billions of years?
- A. A flat surface is gradually pushed up into higher and higher mountains until the Earth is covered with mountains.
 - B. High mountains gradually wear down until most of the Earth is at sea level.
 - C. High mountains gradually wear down as new mountains are continuously being formed, over and over again.
 - D. High mountains and flat plains stay side by side for billions of years with little change.

11. ¿Cuál de las siguientes afirmaciones describe MEJOR los cambios de la superficie de la Tierra a través de miles de millones de años?

- A. Una superficie plana se transforma gradualmente en montañas altas hasta que la Tierra es cubierta con montañas.
- B. Las montañas muy altas fueron bajando gradualmente hasta que la mayor parte de la Tierra bajó al nivel del mar.
- C. Las montañas muy altas fueron transformándose gradualmente en superficies planas que quizás se transformaron otra vez en montañas y así sucesivamente.
- D. Las montañas muy altas y las superficies planas se mantuvieron por miles de millones de años con pequeños cambios.

J8. Which gas could cause a glowing splint to burst into flame?

- A. Neon
- B. Oxygen
- C. Nitrogen
- D. Carbon dioxide

J8. ¿Qué gas puede provocar que un pedazo de carbón al rojo vivo produzca una flama repentinamente?

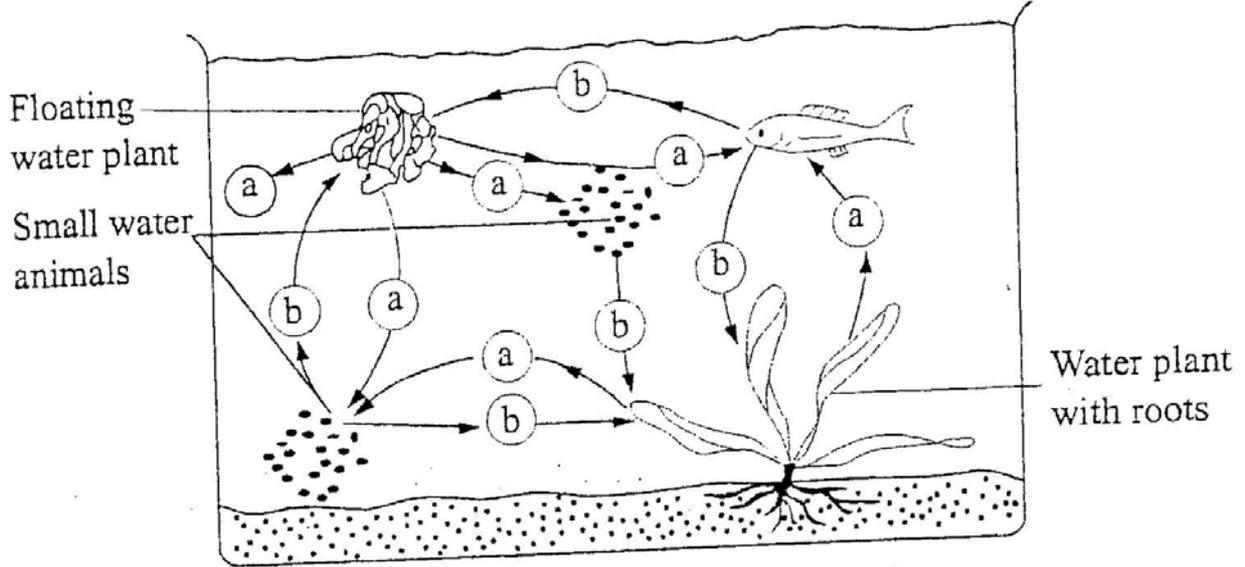
A. Neón.

B. Oxígeno.

C. Nitrógeno.

D. Dióxido de carbono.

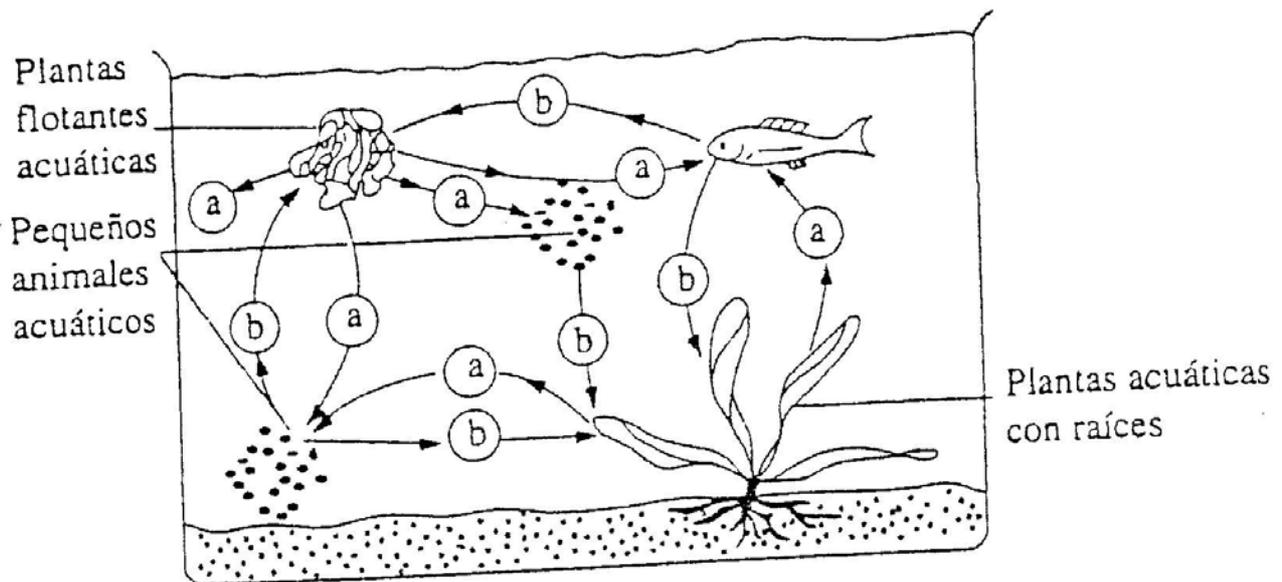
K11. The diagram below shows an example of interdependence among aquatic organisms. During the day the organisms either use up or give off (a) or (b) as shown by the arrows.



Choose the right answer for (a) and (b) from the alternatives given.

- A. (a) is oxygen and (b) is carbon dioxide.
- B. (a) is oxygen and (b) is carbohydrate.
- C. (a) is nitrogen and (b) is carbon dioxide.
- D. (a) is carbon dioxide and (b) is oxygen.
- E. (a) is carbon dioxide and (b) is carbohydrate.

K11. El dibujo muestra un ejemplo de interdependencia entre organismos acuáticos. Durante el día los organismos consumen o liberan (a) o (b) como indican las flechas.

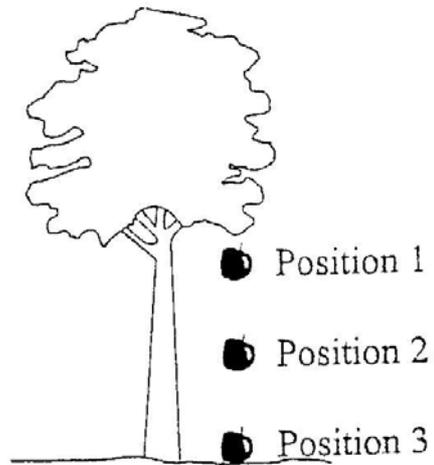


¿Cuáles son las sustancias (a) y (b)?

- A. (a) es oxígeno y (b) es dióxido de carbono.
- B. (a) es oxígeno y (b) es un carbohidrato.
- C. (a) es nitrógeno y (b) es dióxido de carbono.
- D. (a) es dióxido de carbono y (b) es oxígeno.
- E. (a) es dióxido de carbono y (b) es un carbohidrato.

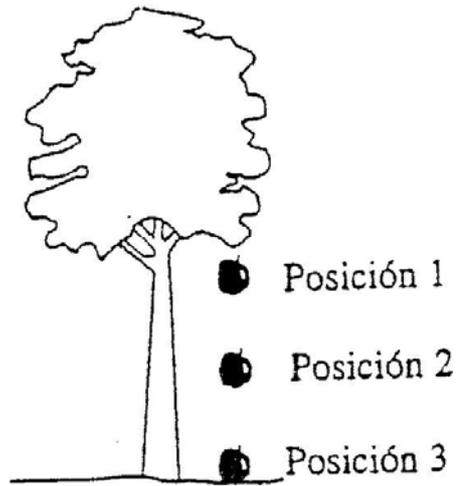
K17. The drawing shows an apple falling to the ground. In which of the three positions does gravity act on the apple?

- A. 2 only
- B. 1 and 2 only
- C. 1 and 3 only
- D. 1, 2, and 3



K17. El dibujo muestra una manzana cayendo al suelo. ¿En cuál de las tres posiciones, la gravedad actúa sobre la manzana?

- A. Sólo en la posición 2.
- B. Sólo en la posición 1 y 2.
- C. Sólo en la posición 1 y 3.
- D. Posiciones 1, 2 y 3.



N4. Years ago farmers found that corn plants grew better if decaying fish were buried near by. What did the decaying fish probably supply to the plants to improve their growth?

- A. energy
- B. minerals
- C. protein
- D. oxygen
- E. water

N4. Hace años los agricultores encontraron que las plantas de maíz crecían mejor si se enterraba pescado en mal estado cerca de éstas. ¿Qué es lo que probablemente aporta el pescado en mal estado que favorece el crecimiento de las plantas?

- A. Energía
- B. Minerales
- C. Proteínas
- D. Oxígeno
- E. Agua

N5. One of the principal causes of acid rain is

- A. waste acid from chemical factories being pumped into rivers
- B. acid from chemical laboratories evaporating into the air
- C. gases from burning coal and oil dissolving in water in the atmosphere
- D. gases from air conditioners and refrigerators escaping into the atmosphere

N5. Una de las principales causas de la lluvia ácida es:

- A. Los desechos ácidos de fábricas de productos químicos que son bombeados a los ríos.
- B. El ácido de laboratorios químicos que se evapora en el aire.
- C. La unión de los gases de la combustión del carbón y el petróleo con el agua que existe en la atmósfera.
- D. El gas de aparatos de aire acondicionado y refrigeradores que escapa hacia la atmósfera.

Q15. Which is NOT an example of a chemical change?

- A. Boiling water
- B. Rusting iron
- C. Burning wood
- D. Baking bread

Q-15

Q15. ¿Cuál de los siguientes NO es un ejemplo de un cambio químico?

- A. Agua hirviendo.
- B. Fierro oxidándose.
- C. Madera quemándose.
- D. Horneando pan.

Q16. How long does it take light from the nearest star other than the Sun to reach Earth?

- A. Less than 1 second
- B. About 1 hour
- C. About 1 month
- D. About 4 years

Q-16

Q16. ¿Cuánto tiempo le toma a la luz de la estrella más cercana (diferente del Sol) llegar a la Tierra?

- A. Menos de 1 segundo.
- B. Cerca de 1 hora.
- C. Cerca de 1 mes.
- D. Cerca de 4 años.