

**VALORACIÓN DE LA EVALUACIÓN DEL CURSO DE INGLÉS A TRAVÉS DE  
EXÁMENES COMPUTARIZADOS**

**ASSESSMENT OF THE EVALUATION OF THE ENGLISH COURSE THROUGH  
COMPUTERIZED TESTS**

Leonora Estrada Flores / Mabel García Mendoza / Esmeralda Irais León  
López / Mónica Loya Rangel / María Sánchez Santiago / Luis Fernando  
González Juárez

**Nota sobre las autoras:**

1. Maestra en Comunicación y Estudios de la Cultura; Formador de Inglés C.

 <https://orcid.org/0000-0003-0321-8104>

2. Maestra en Administración; Formador de Inglés C.

 <https://orcid.org/0000-0002-8091-9468>

3. Licenciada en Enseñanza de Lengua Inglesa; Formador de Inglés C.

 <https://orcid.org/0000-0002-9461-1798>

4. Maestra en Ingeniería Económica y Financiera; Formador de Inglés C.

 <https://orcid.org/0000-0002-1938-4994>

5. Maestra en Desarrollo Educativo; Profesor Investigador Asociado C.

 <https://orcid.org/0000-0002-5764-5091>

6. Maestro en Educación; Profesor Investigador Asociado C.

 <https://orcid.org/0000-0001-9873-2880>

Todas se desempeñan actualmente como docentes-investigadoras en el Centro Regional de Educación Normal “Benito Juárez” de Pachuca, Hidalgo, México.

Esta investigación fue financiada con recursos de las autoras. Las autoras no tienen ningún conflicto de interés al haber hecho esta investigación.

Remita cualquier duda sobre este artículo al siguiente correo electrónico: [leonora.estrada@crenbenitojuarez.edu.mx](mailto:leonora.estrada@crenbenitojuarez.edu.mx)

Recibido: 01/08/2021    Corregido: 10/03/2022    Aceptado: 01/04/2022



Copyright (c) 2022 Leonora Estrada Flores, Mabel García Mendoza, Esmeralda Irais León López, Mónica Loya Rangel, María Sánchez Santiago y Luis Fernando González Juárez. Este texto está protegido por una licencia [Creative Commons 4.0](https://creativecommons.org/licenses/by/4.0/).

Leonora Estrada Flores / Mabel García Mendoza / Esmeralda Irais León López /  
Mónica Loya Rangel / María Sánchez Santiago / Luis Fernando González  
Valoración de la evaluación del curso de inglés  
a través de exámenes computarizados  
Revista *Xihmai* XVII (33), 9-42, enero-junio 2022

## **VALORACIÓN DE LA EVALUACIÓN DEL CURSO DE INGLÉS A TRAVÉS DE EXÁMENES COMPUTARIZADOS**

### **ASSESSMENT OF THE EVALUATION OF THE ENGLISH COURSE THROUGH COMPUTERIZED TESTS**

#### **Resumen**

Ante el confinamiento provocado por la pandemia de COVID-19, algunos docentes comenzaron a utilizar instrumentos de la educación a distancia. Esto incluyó exámenes computarizados, realizados en línea mediante plataformas digitales, calificados de forma automatizada, con el consiguiente cuestionamiento de si la evaluación computarizada podría seguir siendo usada cuando se regrese a las clases presenciales.

Con el fin de comparar los exámenes tradicionales con los exámenes computarizados, en esta investigación se utilizaron el enfoque cualitativo, técnicas de recolección de datos del análisis documental, la encuesta y las escalas estimativas. Los exámenes computarizados se usaron durante todo un semestre (septiembre 2020-enero 2021), procediendo a evaluar su uso al finalizar el mismo.

Al final, se encontró que los exámenes tradicionales son superiores a las evaluaciones en línea. Sin embargo, los exámenes computarizados tienen ventajas que conviene aprovechar, por lo que son una opción interesante como evaluación complementaria en los cursos de inglés.

**Palabras clave:** *evaluación, exámenes computarizados, evaluación de inglés, exámenes en línea.*

#### **Abstract**

Because of the confinement caused by the COVID-19 pandemic, some teachers started using remote education tools. This included computerized examinations, administered online through digital platforms, which are graded in an automated manner, thereby questioning whether computerized evaluation could continue to be used when returning to face-to-face classes.

In order to compare traditional examinations with computerized examinations, this research used the qualitative approach, the research-action method and the data collection techniques of documentary analysis, survey and estimated scales. The

computerized tests were used during a whole semester (September 2020-January 2021), proceeding to assess its use by the end of said semester.

In the end, traditional exams were found to be superior to online assessments. However, computerized exams have advantages that should be leveraged, making them an interesting option as complementary assessment in English courses.

**Keywords:** *evaluation, computerized tests, English assessment, online tests.*

## **I. PLANTEAMIENTO DEL PROBLEMA**

La UNESCO afirma que el confinamiento y el cierre de los centros educativos en América Latina ha afectado a más de 150 millones de estudiantes. Sin embargo, también señala que han sido acciones ineludibles: “estas medidas dolorosas son necesarias y deberán mantenerse hasta que el riesgo disminuya” (Uribe et al., 2020). Muy a pesar de que la comunidad médica trabaja arduamente para proveer a la población con una cura y una vacuna, la fecha en que esa disminución del riesgo ocurrirá sigue siendo incierta.

Los docentes de todo el mundo hemos tenido que adaptar nuestras formas de enseñar, haciendo uso de la tecnología a nuestro alcance, aprendiendo a usar ciertas herramientas en el proceso. En palabras de Audrey Azoulay, directora general de la UNESCO, “este esfuerzo es una oportunidad para repensar la educación, ampliar el aprendizaje a distancia y hacer que los sistemas educativos sean más resistentes, abiertos e innovadores” (citado en Uribe et al., 2020).

Ahora bien, uno de los componentes de mayor importancia en la educación a distancia es la evaluación. Como respuesta ante la contingencia sanitaria, parte de la evaluación del curso de inglés durante el semestre septiembre 2020-enero 2021 en el Centro Regional de Educación Normal (CREN) “Benito Juárez” de Pachuca, Hidalgo, México, fue realizada mediante pruebas en línea a través la plataforma Moodle, que de manera institucional se conoce como Campus Virtual. Dado que, previo a la pandemia, este tipo de herramientas no habían sido usadas, y en virtud de que su implementación ha sacado a relucir ventajas y desventajas, se decidió reflexionar sobre el uso de la evaluación en línea a través de pruebas computarizadas, para así valorar si se pudiese continuar con su uso después del confinamiento.

## **PREGUNTA**

¿Cuáles serían los beneficios de evaluar el curso de inglés mediante exámenes computarizados aplicados a través del Campus Virtual para favorecer el trabajo de los docentes del CREN Benito Juárez de Pachuca, Hidalgo, durante el semestre septiembre 2020- enero 2021?

## **PROPOSICIÓN**

La evaluación del curso de inglés se ve beneficiada cuando se utilizan exámenes computarizados aplicados a través de una plataforma digital ya que son más confiables, tienen un mayor grado de validez en cuanto a su construcción y son más prácticos en lo que respecta al uso de recursos, por lo que presentan más ventajas para los docentes que deseen usarlos, comparados con los exámenes tradicionales en papel.

## **II. MARCO TEÓRICO**

### **2.1. La evaluación y los exámenes**

Ryan et al. plantean la evaluación de los aprendizajes como “un proceso mediante el cual los estudiantes ganan una comprensión de sus propias competencias y progreso, así como un proceso mediante el cual son calificados” (2000, p. 128). Es decir, la evaluación tiene una doble función: por un lado otorga a los estudiantes información sobre su propio aprendizaje, que en los cursos de inglés se expresa en competencias sobre aquello que los estudiantes son capaces o no de realizar (*can do statements*); por otro lado, dota al docente de la información necesaria para establecer si hay contenidos que pueden considerarse ya consolidados, si se deben implementar otras estrategias de enseñanza-aprendizaje, si el estudiante puede continuar en el siguiente nivel o es necesario reforzar ciertos aprendizajes antes de continuar, para asignar una calificación, entre muchas otras acciones.

Castillo y Cabrerizo (2010) describen la evaluación como,

Un proceso o procedimiento que permite no solo saber cómo han ido avanzando los estudiantes en su trayecto de aprendizaje, sino que también es un medio u ocasión

relevante de enseñanza y de aprendizaje, de actuación formativa y, en definitiva, de educación de los alumnos. (p. 372)

Es decir, que no hay que olvidar que la evaluación es también un medio para que los estudiantes aprendan.

Por su parte, Morgan y O'Reilly (2002) conciben “la evaluación como la maquinaria (motor) que dirige y da forma al aprendizaje, más que simplemente un evento final que califica y reporta el desempeño” (p. 13). Estos autores consideran que este enfoque permite encontrar nuevas oportunidades para promover, a través de la evaluación, tipos de aprendizaje más útiles y deseables.

Ahora bien, los autores antes mencionados hacen una distinción entre *evaluación formativa* y *evaluación sumativa*. En cuanto a la evaluación formativa, indican que su propósito es “obtener información acerca del progreso de un estudiante en particular, para darle retroalimentación a ese estudiante y a sus profesores” (p. 49). Para Morgan y O'Reilly (2002), la evaluación formativa “comprende todas aquellas actividades diseñadas para motivar, para aumentar la comprensión y para proporcionar a los estudiantes una indicación de sus progresos” (p. 15). Es importante resaltar que dicha retroalimentación puede ayudar a reforzar la motivación de los estudiantes, así como proporcionar al docente información para mejorar las estrategias de enseñanza-aprendizaje, sobre características de los estudiantes, entre otros.

Respecto a la evaluación sumativa, esta se realiza “para poder dar una calificación o grado que indica cómo se compara el desempeño de un estudiante particular con el de sus compañeros o con un conjunto de criterios” (Ryan et al., 2000, p. 49). Esta descripción coincide con la propuesta por Morgan y O'Reilly, quienes señalan que su principal propósito es “registrar o reportar una estimación de los logros de los estudiantes” (2002, p. 15). Así, la evaluación sumativa frecuentemente se realiza al final del curso, y conduce a calificaciones.

Por otro lado, en los ambientes educativos se realiza una evaluación continua, la cual tiene componentes tanto formativos como sumativos. La evaluación continua tiene las siguientes características (Morgan y O'Reilly, 2002, p. 16):

- Proporciona alguna estructura al aprendizaje.
- Descompone la carga de evaluación en partes manejables.
- Es alentadora, motivante, crea confianza.
- Proporciona una fuente de diálogo favorable entre profesores y alumnos.
- Facilita a los alumnos una visión de sus progresos, incluyendo el desarrollo de su comprensión y dominio de competencias.

Por lo tanto, se puede concluir que el objetivo principal de la evaluación es la de proporcionar información sobre el desarrollo o progreso de cada estudiante, información a partir de la cual tanto estudiantes como docentes pueden tomar decisiones.

Cabe señalar que el dominio de una lengua tiene sus propias características y tipos de evaluación. Jeremy Harmer numera 5 tipos de evaluación o exámenes (2014, p. 166):

- a) **Exámenes de ubicación.** Estos suelen ser usados al inicio de un curso o cuando el estudiante ingresa a la institución, con el fin de ubicarlo en el nivel de dominio correspondiente.
- b) **Exámenes de progreso.** Son utilizados a lo largo del curso para comprobar el avance de los estudiantes y cómo han asimilado lo estudiado durante la semana, la quincena o el mes. En otras palabras, sirven para que los estudiantes muestren lo que saben y pueden hacer, o bien, lo que no saben y no pueden hacer.
- c) **Exámenes de logro o de salida.** Su finalidad es evaluar el total de los contenidos del curso y comúnmente son usados para decidir si el estudiante aprueba o no el curso, o bien, avanza o no al siguiente nivel.
- d) **Exámenes de dominio o competencia.** Según Harmer, estos exámenes están diseñados para mostrar qué nivel ha alcanzado un estudiante en cualquier momento, y son utilizados por empleadores y universidades, por ejemplo, que quieren una medida confiable de las habilidades de idioma de un estudiante.
- e) **Evaluación continua.** Es el tipo de evaluación que mide el avance de los estudiantes mientras está sucediendo; los logros de los estudiantes se observan durante todo el periodo de aprendizaje, y no solamente al final. Esta evaluación incluye, por ejemplo, los portafolios de evidencias.

Por otro lado, en cuanto al tipo de reactivos en los exámenes, existen diferentes clasificaciones. Harmer (2014) indica dos tipos principales: discretos y uso integrador del lenguaje: “Las pruebas de reactivos discretos solo evalúan una

cosa a la vez” (p. 168), por ejemplo, evaluar un tiempo verbal o una palabra, “mientras que las pruebas integradoras consisten en pedir a los estudiantes que utilicen una variedad de habilidades del lenguaje para completar una tarea con éxito” (p. 168). Adicionalmente, el mismo autor hace una distinción entre ítems directos e ítems indirectos: “Un reactivo directo es aquel que pide a los alumnos que hagan algo con el lenguaje” (p. 168), por ejemplo, escribir una carta, leer y responder a un artículo de periódico, participar en una conversación, entre otros. Mientras que los reactivos indirectos tienden a ser integradores: “Los elementos de prueba indirectos son aquellos que ponen a prueba el conocimiento del lenguaje de los estudiantes en lugar de conseguir que lo usen” (p. 168). Algunos ejemplos de elementos de prueba indirectos son las preguntas de opción múltiple, las preguntas de rellenar (con una palabra o una frase), los ejercicios de relacionar columnas, los ejercicios en que se deben ordenar palabras para formar una oración, entre otros.

Esta clasificación puede relacionarse con la distinción hecha por Joseph E. Aoun (2018) entre pensamiento convergente y pensamiento divergente. Aoun afirma que “Cuando una persona emplea el pensamiento convergente, se enfoca en encontrar la respuesta única y ‘correcta’ a un problema o tarea” (p. 49). El autor explica que el pensamiento convergente consiste en sopesar datos y alternativas para lograr el mejor resultado, un resultado que solo puede ser correcto o incorrecto. Responder preguntas en un examen de opción múltiple es el mejor ejemplo de este tipo de pensamiento. Además, el autor señala que este tipo actividad mental es justamente aquella en la que las computadoras y la inteligencia artificial están mejorando.

En cambio, el pensamiento divergente “es la generación creativa de respuestas múltiples en un libre flujo de ideas” (p. 50). Aoun señala que el pensamiento divergente se relaciona con el juego, la curiosidad y la toma de riesgos, y sus ejemplos son la lluvia de ideas y la escritura libre, sin atención a la estructura o la gramática.

Por su parte, Frodden et al. (2004) distingue dos tipos de instrumentos de evaluación: los duros y los suaves:

«Duro» se refiere a una forma tradicional y a un propósito de evaluar que hace hincapié en la objetividad, precisión, fiabilidad, y se centra en el producto en lugar

de en el proceso. Utiliza principalmente datos cuantitativos proporcionados por instrumentos como exámenes y pruebas. (p. 8)

Como puede observarse, los términos ya descritos de reactivos discretos, ítems indirectos, y pensamiento convergente tienen relación con este tipo de evaluación. En cambio, la evaluación «suave», más relacionada con los ítems directos de Harmer, es definida como “una forma naturalista y alternativa y a la finalidad de evaluar, [...] un tipo de evaluación que, al ser continua, tiene en cuenta el proceso de aprendizaje del estudiante para promoverlo” (Frodden et al., 2004, p. 8). Esta descripción encaja con las de las pruebas integradoras, ítems directos y promoción del pensamiento divergente.

## **2.2. Cualidades de una prueba o examen**

En su *Análisis de los instrumentos de evaluación utilizados en la enseñanza de lenguas extranjeras* (2004), Frodden y sus colegas recuperan las seis cualidades propuestas por Bachman y Palmer que se deben considerar en el diseño de un examen: (1) confiabilidad, (2) validez de la construcción, (3) autenticidad, (4) interactividad, (5) impacto y (6) practicidad. Las descripciones de cada cualidad son las siguientes:

### **a) Fiabilidad**

Esta cualidad se relaciona con la consistencia de la medición. Si alguien realiza la misma prueba en dos ocasiones diferentes, debe tener resultados similares. Además, si dos o más examinadores dan la misma o similar puntuación a una prueba, podemos decir que la prueba es fiable (Frodden et al., 2004). Un examen poco confiable, entonces, sería aquel que no muestre consistencia en sus resultados, cuando este se lleva a cabo en circunstancias similares.

Esta descripción coincide con la hecha por Jeremy Harmer (2014), quien afirma que un examen es confiable cuando no solamente es fácil de calificar, sino que “cualquiera que lo califique debe obtener el mismo resultado que alguien más” (p. 167). Respecto a esto, el autor nos recuerda que cada docente tiene su propio estilo y criterios al calificar un examen, por lo que las evaluaciones deben ser diseñadas para minimizar el efecto de los estilos individuales de calificación. En palabras del autor: “siempre existirá el peligro de que cuando las pruebas impliquen algo que no sea preguntas calificadas por

computadora, diferentes resultados serán dados por diferentes evaluadores” (p. 167). Así, los instrumentos de evaluación «duros» tienen mayor fiabilidad que los instrumentos «suaves», sobre todo si son calificados por un software en lugar de un ser humano.

#### **b) Validez de la construcción**

Se refiere a los conocimientos o habilidades que la prueba pretende evaluar (Frodden et al., 2004); estos pueden ser estructurales (gramática y vocabulario) o comunicativos (*reading, listening, speaking & writing*). Depende del diseño de la prueba y, siendo así, los instrumentos «duros», con una sola respuesta correcta, tienen menor validez que aquellos basados en reactivos «suaves», con preguntas abiertas, exposiciones, entre otros.

#### **c) Autenticidad**

Se refiere al “grado de correspondencia de las características de una determinada tarea en un examen de idioma con las características de una tarea de uso del lenguaje objetivo” (Frodden et al., 2004, p. 15). Es decir, qué tan cercano es el lenguaje usado en el examen con el lenguaje que el estudiante usaría en una situación de la vida real. Dado que es difícil inferir el desempeño de un estudiante en una situación real a través de los instrumentos «duros» de evaluación, se puede considerar que estos tienen menor autenticidad que los instrumentos «suaves».

#### **d) Interactividad**

Según Frodden y sus colegas (2004), se entiende por interactividad a la forma en que la capacidad del lenguaje, el conocimiento temático y la *esquemata* afectiva de quien toma el examen se ponen a prueba. Un examen se considera interactivo cuando el estudiante utiliza un conjunto de estrategias metacognitivas, decide qué va a hacer, analiza lo que necesita para resolver la tarea de la prueba y las posibilidades que tiene para completarla satisfactoriamente, y decide cómo utilizará el conocimiento que posee.

El conocimiento temático y la *esquemata* afectiva están directamente ligados con la cultura general e intereses de los estudiantes, es decir, los temas que conocen y les son interesantes, agradables, sensibles, etc. Al ser una cuestión subjetiva, la interactividad es difícil de medir. Sin embargo, se puede afirmar

que los instrumentos «duros» tienen menor interactividad que instrumentos «suaves» tales como los proyectos o la redacción de textos.

#### **e) Impacto**

El impacto de una prueba se refiere a la forma en que sus resultados afectan a individuos, instituciones o sociedad. El conocimiento de temas, la familiaridad con las técnicas e instrucciones del examen, el *input*, la retroalimentación y las decisiones tomadas con base en los resultados del examen son algunos ejemplos de cómo las puntuaciones de las pruebas podrían afectar a las personas positiva o negativamente (Frodden et al., 2004).

Los puntajes de las pruebas proporcionan a los estudiantes comentarios sobre su rendimiento y promueven la introspección sobre las estrategias de aprendizaje que están utilizando, por ejemplo. Además, los estudiantes cuyos profesores proporcionan comentarios más descriptivos sobre los resultados de sus pruebas pueden ser más capaces de introspección y mejorar. En este sentido, no existe mucha distinción entre instrumentos de evaluación «suaves» y «duros», ya que el impacto de sus resultados depende de la forma en que el docente transmite la retroalimentación a los estudiantes y la manera en que cada uno reacciona ante ella. Sin embargo, si los resultados de una prueba computarizada solo son conocidos a través de la plataforma a través de la cual esta se aplica, sin más información que la calificación obtenida, se puede afirmar que los exámenes computarizados tienen menor impacto que los exámenes en papel, ya que estos últimos muchas veces son devueltos al estudiante con las debidas correcciones hechas por el docente.

#### **f) Practicidad**

Es la relación entre los recursos que se requerirán en el diseño, desarrollo y uso de la prueba y los recursos que estarán disponibles para estas actividades (Frodden et al., 2004). Esto significa que los recursos humanos, materiales y de tiempo (considerando elaboración, implementación y revisión), así como la posibilidad de realizar la prueba en diferentes lugares y momentos, son esenciales para estimar la practicidad de una prueba.

Ahora bien, los instrumentos de evaluación «duros» frecuentemente implican una prueba escrita objetiva, que podría considerarse fácil y rápida de elaborar y calificar por parte del docente. Por su parte los recursos necesarios para

evaluar mediante instrumentos de evaluación «suaves» dependerán de las decisiones tomadas tanto por el docente como por los estudiantes. Este tipo de instrumentos abarca un amplio abanico de posibilidades, ya que pueden ir desde un juego de rol breve (con alto grado de practicidad) hasta la elaboración de un proyecto durante todo el curso (bajo grado de practicidad). Sin embargo, en términos generales, se puede considerar que los instrumentos «duros» tienen mayor grado de practicidad que los instrumentos «suaves».

### **2.3. Evaluación en línea**

Quesada define la enseñanza a distancia como aquella “que se realiza fuera de los recintos escolares, a partir de recursos como la internet, páginas web, foros de discusión, chat, correo electrónico y también videoconferencia, audio, video; aunque puede incluir algunas actividades presenciales” (Quesada, 2006, p. 2). Debido a la contingencia sanitaria, las clases impartidas en el CREN “Benito Juárez” tuvieron que mudarse a la modalidad a distancia. En el periodo estudiado, las sesiones de videoconferencia se realizaron utilizando herramientas tales como Google Meet y Zoom, mientras que los documentos, evaluaciones y tareas se gestionaron a través del Campus Virtual.

Acerca de la educación en línea de la que los sistemas educativos de todo el mundo han hecho uso a partir de la contingencia sanitaria, la Inspección Técnica de Uruguay indica que: “No es cuestión de transferir toda la responsabilidad educativa a la familia ni pretender que la tecnología *per se* cumpla el rol de enseñante” (Consejo de Educación Inicial y Primaria [CEIP], 2020, p. 2). Está claro que la tecnología por sí misma no suple al docente, ni que esa tarea deba o pueda ser realizada por los padres o tutores. Es más, en el caso de los estudiantes del CREN “Benito Juárez”, como el resto de los estudiantes de nivel superior, la responsabilidad educativa recae en sí mismos y les lleva a ser autodidactas. El texto continúa: “es tiempo de generar una nueva ‘instalación’ educativa que exige de cada docente habitar otras formas y otros modos, siempre con la misma pasión para transformar el saber en conocimiento” (CEIP, 2020, p. 2). De esta forma, la evaluación también tiene que ser adaptada a los contextos virtuales.

La evaluación en sí misma y la educación a distancia tienen cada una sus propias complejidades. Por tanto, la evaluación a distancia se vuelve doblemente compleja. Rodríguez (2014) señala que la evaluación a distancia

está planteada como un problema aún no resuelto, y que el fortalecimiento de estrategias de evaluación para este contexto particular es una necesidad. Así, aunque no se trata de un estudio exhaustivo sobre la evaluación en línea, en los siguientes apartados se describen las características de la evaluación de la materia de inglés en el CREN “Benito Juárez” previo y durante la contingencia sanitaria, y se realiza una comparación de ambos tipos de evaluación según una serie de cualidades que deben tener las evaluaciones.

### **III. METODOLOGÍA**

#### **3.1. Enfoque de la investigación**

Para esta investigación, se eligió el enfoque cualitativo, ya que está orientado a comprender los fenómenos, examinándolos desde el punto de vista de los participantes en un ambiente natural y en relación con su contexto (Hernández et al., 2014). En este caso, los participantes son los docentes del CREN “Benito Juárez” cuando hacen uso de los exámenes computarizados como instrumento para evaluar el curso de inglés.

#### **3.2. Técnicas e instrumentos**

Las principales técnicas de recolección de datos utilizadas en esta investigación fueron el análisis documental acerca de los tipos de evaluación, la evaluación del curso de inglés y el uso de pruebas computarizadas como instrumento de evaluación; la encuesta y, finalmente, las escalas estimativas para valorar los exámenes tradicionales en papel comparados con los exámenes computarizados y en línea.

#### **3.3. Población**

En virtud de que esta investigación está orientada a encontrar los beneficios que tiene el uso de los exámenes computarizados para favorecer el trabajo docente, la población de este estudio abarcó a la totalidad de los docentes que impartieron el curso de inglés en el CREN “Benito Juárez” de Pachuca, Hidalgo, durante el semestre septiembre 2020-enero 2021, es decir, seis docentes. No fue necesario, por tanto, establecer una muestra, ya que las

observaciones y reflexiones de los seis docentes pudieron ser incluidos en el estudio.

### **3.4. Proceso metodológico seguido**

Para la realización de esta investigación se desarrollaron las siguientes actividades:

- Sistematización de información para conocer el estado del arte a partir de fuentes de consulta.
- Revisión de bibliografía sobre tipos de evaluación, evaluación del curso de inglés, y evaluación a través de plataformas digitales y pruebas computarizadas.
- Descripción del proceso de evaluación previo a la pandemia.
- Observación y registro del proceso de evaluación llevado a cabo durante el semestre septiembre 2020-enero 2021.
- Selección y propuesta sobre los indicadores requeridos para valorar los instrumentos de evaluación seleccionados.
- Obtención de datos y reflexiones a través de encuestas a docentes de inglés.

## **IV. DESARROLLO Y DISCUSIÓN**

### **4.1. Evaluación de los cursos de inglés en el CREN “Benito Juárez”**

La evaluación de cada curso de inglés queda establecida desde la planeación. En dicho documento, se puede observar que la evaluación de los cursos de inglés en el CREN “Benito Juárez” promueve tanto la evaluación sumativa como la evaluación formativa. La primera es realizada mediante exámenes de medio y fin de curso, mientras que la segunda se hace a través de otras herramientas, como portafolio de evidencias y evaluación continua durante clase. La aplicación de exámenes es vista como un componente de la evaluación, sin embargo, no debe tener la mayor parte del peso en la decisión de acreditar o no el curso, dado que la evaluación se concibe como parte de la cotidianeidad en el aula. Así mismo, se invita a implementar la autoevaluación a fin de crear conciencia en el estudiante de sus propios errores y facilitar la autocorrección.

Por otro lado, en la planeación también se indica que las evidencias se dividen en tres tipos:

- a) Evidencias de conocimiento, mediante las cuales el estudiante normalista demuestra el dominio de un mínimo de elementos teóricos básicos con relación al idioma inglés, para ello aplican cuestionarios, preguntas, reflexiones y análisis de lo trabajado en cada unidad.
- b) Evidencias de producto, el estudiante realiza *quizzes* (pruebas rápidas) al final de cada unidad para evaluar su dominio de la lengua. También, se incluyen presentaciones de acuerdo con los temas trabajados durante la unidad, videos, escritos y elaboración de audios, entre otros.
- c) Evidencias de desempeño se manifiestan al participar en debates, demostraciones prácticas y exposiciones, seleccionado al alumno que expondrá en el momento de la clase, así como las actitudes en todo momento.

Previo a la pandemia provocada por el virus de COVID-19, los cursos de inglés en el CREN “Benito Juárez” se evaluaba bajo los siguientes porcentajes y criterios:

	RUBRO	PORCENTAJE
Evaluación Sumativa	Examen ( <i>Grammar &amp; Vocabulary, Reading, Listening, Writing y Speaking</i> )	60%
	Trabajo y participación en clase	15%
Evaluación Formativa	Tareas	10%
	<i>Quiz</i> al final de cada unidad	10%
Evaluación Iluminativa	Valores (honestidad, respeto y responsabilidad)	5%
TOTAL		100%

Tabla 1: Porcentajes de evaluación previos a la contingencia sanitaria. Elaboración propia.

Ahora bien, en cuanto a la clasificación de tipos de evaluación hecha por Harmer (2014), descrita en el apartado 2.1 de esta investigación (II. Marco teórico – La evaluación y los exámenes), en el CREN “Benito Juárez” se han realizado los siguientes tipos de evaluación: exámenes de ubicación, exámenes de dominio o competencia, evaluación continua, exámenes de progreso y exámenes de logro o de salida. A continuación, se describen únicamente los

dos últimos, ya que son los únicos que se han hecho también a través de exámenes computarizados.

- a) **Exámenes de progreso.** Se aplica un *quiz* al final de cada unidad para comprobar el avance de los estudiantes de manera continua. Previo a la pandemia, dicho *quiz* podía ser un dictado, un cuestionario breve o algún ejercicio similar. Durante las clases a distancia, los *quizzes* consistieron en pruebas realizados a través del Campus Virtual (Moodle), siendo en su mayoría ejercicios de opción múltiple que evaluaban temas de gramática y vocabulario.
- b) **Exámenes de logro o de salida.** Se aplica un examen a medio semestre (examen parcial) y otro al final de este (examen final), los cuales evalúan el total del contenido visto durante el curso, por lo que cuentan con ejercicios de gramática y vocabulario, así como *reading*, *listening* y *writing*, mientras que la habilidad de *speaking* se evalúa a través de entrevistas breves con los alumnos de manera individual, por parejas o a través de exposiciones individuales o en equipo con los estudiantes más avanzados. No obstante, la decisión de si el estudiante avanza o no de nivel es considerado junto con otros rubros, ya que este se promedia con otros aspectos como participación en clase, tareas, etc.

Ahora bien, antes de la contingencia sanitaria, los exámenes parcial y final eran aplicados en papel y de manera presencial. Durante el semestre septiembre 2020-enero 2021, los exámenes se realizaron a través del Campus Virtual, siendo en su mayoría reactivos de opción múltiple, es decir, pruebas de reactivos discretos, esto es, reactivos «duros». Las habilidades de *writing* y *speaking* se evaluaron a través de otras herramientas.

#### **4.2. Evaluación de los cursos de inglés en el CREN durante el semestre septiembre 2020-enero 2021**

Al inicio del confinamiento, a partir del 23 de marzo de 2020, los docentes tuvieron que adaptar su trabajo para poder usar recursos en línea y dar clases a distancia. El CREN “Benito Juárez” comenzó a usar el Campus Virtual a inicios del mes de abril. Dicha plataforma es un Moodle, y cuenta con múltiples

recursos para la educación a distancia. Su implementación requirió un proceso de aprendizaje por parte de estudiantes y docentes.

Gracias a la experiencia ganada durante el cierre del semestre enero-junio 2020, los diferentes cursos sobre educación a distancia que organizó la SEP, y ya con mayor conocimiento de las opciones que brinda la plataforma virtual, los docentes de inglés decidieron que los porcentajes de evaluación para el semestre septiembre 2020-enero 2021 serían los siguientes:

	RUBRO	PORCENTAJE
Evaluación Sumativa	Examen ( <i>Grammar and Vocabulary, Reading, Listening, Writing &amp; Speaking</i> )	30%
Evaluación Formativa	Plataforma virtual de Cambridge	45%
	<i>Quiz</i> al final de cada unidad	10%
	Evidencias varias	15%
TOTAL		100%

Tabla 2: Porcentajes de evaluación durante el semestre septiembre 2020-enero 2021. Elaboración propia.

Como se puede observar, se dieron los siguientes cambios:

- Se eliminó el rubro de valores, ya que este se evaluaba durante las clases presenciales y en eventos institucionales.
- Se redujo el valor del examen, ya que en adelante se realizaría en línea. Las características de dichos exámenes se describirán en el apartado 4.4.
- Se eliminó el porcentaje de trabajo y participación en clase. Esta decisión se tomó considerando los posibles problemas de conexión que pudieran tener los estudiantes, lo cual podría afectar su nivel de participación en clase y, por tanto, su calificación.
- Se agregó el trabajo en la plataforma virtual de Cambridge, que para algunos grupos correspondió a la serie de *Touchstone*, mientras que para otros fue la de la serie *Interchange*. En semestres anteriores, la realización de las actividades en la plataforma de Cambridge era considerado dentro del rubro de tareas, y solamente se había trabajado con la serie *Interchange*. Cabe mencionar que se eligió trabajar con este tipo de plataformas ya que ofrecen una gran variedad de actividades donde se ejercitan las cuatro habilidades del idioma y, al

ser plataformas interactivas, resultan entretenidas y atractivas para los estudiantes.

- Se mantuvo el porcentaje de los *quizzes*, pero se cambió de modalidad. Antes de la pandemia, los docentes aplicaban pequeños cuestionarios en papel, *quizzes* de verbos, dictados, ejercicios de *listening* o incluso solicitaban a los estudiantes presentar diálogos en parejas o equipos. Durante el semestre septiembre 2020-enero 2021, los *quizzes* consistieron en pequeñas pruebas en línea donde se evaluó primordialmente gramática y vocabulario. Las características detalladas de estos *quizzes* se describen en el apartado 4.4.
- El rubro de tareas se sustituyó por el de evidencias y se aumentó en un 5%. Las evidencias fueron enviadas de manera digital a los docentes usando principalmente el Campus Virtual. Sin embargo, también se utilizaron otras vías de comunicación disponibles, tales como correo electrónico o servicios de mensajería como Whatsapp y Telegram, o bien, por Dropbox o Drive. Dichas evidencias incluyeron: redacción de oraciones o preguntas aisladas, redacción de textos cortos o párrafos, respuestas personales a una lista de preguntas, elaboración de organizadores tales como diagramas, mapas mentales, etc., fotografías del libro o de hojas de ejercicios (*worksheets*) contestadas, *quizzes* de verbos, envío de audios o videos, salas de chat, entre otros.

#### **4.3. Uso de los exámenes computarizados a través de una plataforma digital como opción ante el confinamiento**

El traslado de la educación presencial a la modalidad a distancia en su momento fue forzado debido a la contingencia sanitaria, y con ello se presentaron diversas dificultades, siendo una de ellas, especialmente, la implementación de la evaluación a distancia a través los entornos virtuales.

Durante la fase exploratoria de esta investigación, se convocó vía Facebook a los docentes de inglés de otras Escuelas Normales a contestar una encuesta sobre su percepción acerca del uso de plataformas digitales para evaluar el curso. Dieciocho docentes contestaron la encuesta, cuyos principales resultados fueron los que se enlistan a continuación:

- 90% de los docentes usaron plataformas digitales para evaluar a sus estudiantes.
- Las principales plataformas usadas fueron Google Classroom y Moodle.
- La mayoría de los docentes (78%) usó preguntas de opción múltiple en sus evaluaciones a través de plataformas digitales.
- Los aspectos por evaluar a través de plataforma fueron: gramática, vocabulario y *listening* en 100%, *reading* en 89%, *writing* en 61% y *speaking* en 44%. Esto es congruente con el hecho de que los aspectos estructurales de la lengua y las habilidades receptivas (*listening* y *reading*) son más factibles de ser evaluados mediante reactivos discretos e indirectos, mientras que las habilidades productivas (*writing* y *speaking*), relacionadas con el uso integrador del lenguaje, necesitan otro tipo de evaluación.
- Respecto a la confiabilidad de los exámenes en línea, el 44% de los docentes considera que son confiables, aunque no totalmente; un 33% piensa que son medianamente confiables; el 17% son totalmente confiables y solo el 6% considera que no para nada confiables.
- En cuanto a las principales ventajas del uso de plataformas para la evaluación, la mayoría de los docentes mencionó en primer lugar el no tener que calificar los exámenes, ya que esto se realiza de manera automática, seguida de la posibilidad de hacer bancos de exámenes para su uso posterior y el menor gasto de papel.
- Las desventajas más mencionadas fueron: problemas de conexión de los estudiantes y la posibilidad de que los estudiantes hicieran trampa.
- Finalmente, el 81% de los encuestados dijo que continuaría usando plataformas virtuales como parte de su evaluación después de la pandemia, mientras que el 19% restante todavía lo estaba considerando.

Los resultados completos de la encuesta pueden consultarse en el Anexo 1, al final de este documento.

#### **4.4. Descripción de la evaluación en línea con exámenes computarizados durante el semestre septiembre 2020-enero 2021: beneficios y retos**

Como ya se describió, la evaluación de los cursos de inglés durante el semestre septiembre 2020-enero 2021 comprendió una evaluación sumativa consistente en un examen parcial y un examen final, y una evaluación formativa consistente en los resultados de la plataforma virtual de Cambridge, la aplicación de un *quiz* al final de cada unidad y evidencias.

Cabe recordar que, antes del confinamiento se aplicaban todos los tipos de evaluación indicados por Harmer (exámenes de ubicación, exámenes de progreso, exámenes de logro o de salida, exámenes de dominio o competencia, evaluación continua) y durante la pandemia solamente se suspendió la aplicación de exámenes de ubicación. Ahora bien, dado que la intención de esta investigación es valorar los beneficios del uso de exámenes computarizados aplicados mediante una plataforma digital durante el semestre mencionado, en adelante el estudio se centrará solamente en los exámenes de progreso (*quizzes*) y los exámenes de salida (parcial y final).

Durante el periodo estudiado, los *quizzes* o exámenes de progreso consistían en pruebas rápidas de entre 10 y 20 reactivos, enfocadas a evaluar primordialmente aspectos estructurales del idioma, es decir, gramática y vocabulario. Los estudiantes contaban con entre 15 y 20 minutos para contestarlos, dependiendo del nivel en que se encontraran, y la presentación de resultados se diseñó para que solo pudieran ser revisados cuando la actividad se consideraba finalizada en la plataforma. Es decir, si el *quiz* era asignado para ser contestado entre las 10:00 y 10:30 am, pero algún estudiante con mayor habilidad lo terminaba a las 10:10, dicho estudiante no podría consultar su resultado hasta las 10:30. Además de poder ver su resultado en términos cuantitativos, cada estudiante podía también revisar cada ítem para identificar en qué habían consistido sus errores.

Los exámenes parcial y final o de salida, por su parte, se trataron de pruebas más largas. Estos exámenes tuvieron entre 30 y 40 reactivos y tuvieron una hora de duración. Al igual que los *quizzes*, los resultados solo podían ser revisados por los estudiantes hasta que la actividad se diera por finalizada en

la plataforma. De igual forma, cada estudiante podía revisar cada uno de los reactivos para conocer sus errores y la razón de su calificación.

Además de evaluar gramática y vocabulario, se incluyeron reactivos de las habilidades receptivas: *listening* y *reading*. Las habilidades productivas (*writing* y *speaking*) fueron evaluadas a través de otros instrumentos.

De manera general, los ítems utilizados tanto en los *quizzes* como en los exámenes parcial y final se pueden clasificar como reactivos discretos e indirectos, que evalúan el pensamiento convergente, ya que se trataban de ejercicios de opción múltiple, ordenar palabras, relacionar columnas, falso/verdadero, rellenar, etc. Este tipo de reactivos tenían una sola respuesta correcta, y estaban enfocados a evaluar los aprendizajes sobre la estructura del lenguaje (gramática y vocabulario), o bien las habilidades de *reading* y *listening*. Son, por lo tanto, instrumentos duros de evaluación.

El uso integrador del lenguaje, con actividades enfocadas al uso del pensamiento divergente, se evaluó a través de las diferentes actividades y evidencias que cada docente fue pidiendo a sus estudiantes a lo largo del semestre, por ejemplo: respuestas individuales o personales a preguntas propuestas por el docente, redacción de textos cortos, envío de audios, entre otros. Esto se evalúa a través de ítems directos, es decir, mediante instrumentos suaves de evaluación.

Ahora bien, el interés primordial de esta investigación es el de determinar los beneficios de evaluar los cursos de inglés mediante exámenes computarizados aplicados a través del Campus Virtual para favorecer el trabajo de los docentes del CREN Benito Juárez de Pachuca, Hidalgo, durante el semestre septiembre 2020-enero 2021. Es por ello por lo que a continuación se utilizan las seis cualidades propuestas por Frodden et al. (2004) para comparar tres tipos de exámenes: (1) exámenes en papel o tradicionales, (2) exámenes computarizados y (3) exámenes en línea (es decir, exámenes computarizados aplicados a distancia).

Asimismo, como se hizo patente gracias a la encuesta descrita en el apartado 4.3, la posibilidad de que los estudiantes pudieran hacer trampa al realizar los exámenes en línea fue una de las mayores preocupaciones de los docentes. Es por ello por lo que esta investigación propone incluir el “nivel de supervisión” como una séptima cualidad, siendo que una prueba tiene un nivel alto de

supervisión cuando la posibilidad de hacer trampa por parte de los estudiantes es baja. En el caso de los exámenes computarizados, se les dividió en dos para evaluar este aspecto: (a) exámenes a distancia, que los estudiantes contestan en casa, y (b) exámenes aplicados en la institución, que es como podrían aplicarse una vez terminado el confinamiento.

Se asignó un puntaje de entre 1 y 5 puntos por rubro, siendo 1 el más bajo, 3 para nivel medio y 5 nivel alto. Asimismo, se dividió la cualidad de “Validez de la construcción” en cuatro: validez al evaluar aspectos estructurales (gramática y vocabulario), validez al evaluar *reading*, al evaluar *listening*, y al evaluar *writing*. No se incluyó “validez al evaluar *speaking*” ya que dicha habilidad no se evaluó a través de instrumentos computarizados. Después de que los seis docentes que impartieron inglés durante el periodo estudiado valoraron cada cualidad según el tipo de examen, se obtuvieron los siguientes resultados:

CUALIDAD	Exámenes tradicionales	Exámenes computarizados		DIFERENCIA
		A distancia (en casa)	Aplicados en la institución	
<b>1. Fiabilidad</b>	3.7	4.7		<b>1 punto</b>
<b>2. Validez de la construcción (promedio)</b>	<b>(4.075)</b>	(3.525)		(0.55 puntos)
a) Validez al evaluar aspectos estructurales	<b>4.0</b>	3.7		0.3 puntos
b) Validez al evaluar <i>reading</i>	<b>4.0</b>	3.7		0.3 puntos
c) Validez al evaluar <i>listening</i>	<b>4.0</b>	3.7		0.3 puntos
d) Validez al evaluar <i>writing</i>	<b>4.3</b>	3.0		<b>1.3 puntos</b>
<b>3. Autenticidad</b>	<b>3.0</b>	2.7		0.3 puntos
<b>4. Interactividad</b>	2.3	<b>2.7</b>		0.4 puntos
<b>5. Impacto</b>	<b>4.0</b>	2.0		<b>2 puntos</b>
<b>6. Practicidad</b>	1.7	<b>4.7</b>		<b>3 puntos</b>
<b>7. Supervisión</b>	<b>4.3</b>	1.0	4.0	<b>3.3</b> y 0.3 <b>puntos</b>
<b>TOTAL</b>	<b>35.3</b>	<b>31.7</b>	<b>34.7</b>	<b>3.6</b> y 0.6 <b>puntos</b>

Tabla 3: Valoración de los tres tipos de exámenes según las cualidades propuestas. Elaboración propia.

Como se puede observar, la diferencia entre el puntaje obtenido por los exámenes tradicionales y los exámenes computarizados es menor a un punto en la mayoría de las cualidades incluidas en validez de la construcción. Asimismo, la diferencia es poca en los criterios de autenticidad e interactividad. Por ello, en los siguientes párrafos solo se discutirán aquellas cualidades en las que la diferencia es amplia, es decir, de un punto o más.

En primer lugar, se observa que los docentes del CREN “Benito Juárez” consideran que las pruebas en línea y computarizadas tienen mayor fiabilidad que las tradicionales. Los docentes solo se encargan de diseñarlas y transferirlas a la plataforma, mas no de calificarlas, ya que esa acción se lleva a cabo de manera automática a través de un software, sin subjetividades ni variaciones.

Respecto al criterio de validez de la construcción, la única diferencia importante está en la validez de los reactivos de *writing*, cuyo puntaje es mayor en los exámenes tradicionales.

Por su parte, se puede observar que los docentes consideraron que existe mucha diferencia en el nivel de impacto de cada examen. Aunque los estudiantes pueden revisar uno a uno los errores en un examen computarizado, se considera que el nivel de retroalimentación es mayor cuando esta se da de manera presencial y personal.

Respecto a la practicidad, es claro que los exámenes computarizados presentan una gran ventaja. En este caso, cabe mencionar tres recursos primordiales: el papel, el tiempo y la ubicación. En cuestión del papel, el uso de exámenes y *quizzes* computarizados durante el periodo estudiado, para la asignatura de inglés, resultó en un ahorro de alrededor de 8,000 hojas en un semestre, lo que potencialmente representa 16 mil al año. Ahora bien, según el sitio [aprender.org](https://aprender.org) (2020), a partir de un solo árbol, se pueden producir aproximadamente 12,000 hojas de papel. Por tanto, se podría afirmar que la implementación de exámenes tradicionales utiliza entre uno y dos árboles por ciclo escolar, mientras que la de los exámenes computarizados y en línea no usan papel en absoluto, siendo claramente más sustentables en cuanto a este recurso se refiere.

En cuanto al recurso del tiempo, hay que considerar las tres fases: elaboración, implementación y revisión. La elaboración de los reactivos de los exámenes tradicionales toma el mismo tiempo que la de los exámenes computarizados y en línea. Sin embargo, los exámenes tradicionales no implican su carga posterior a una plataforma digital, por lo que podría considerarse que son más prácticos en su fase de elaboración. No obstante, también debe tomarse en cuenta el tiempo de fotocopiado y organización para su aplicación. Además, una vez que se cuenta con bancos de preguntas, tanto el tiempo de elaboración como el de implementación de los exámenes computarizados se reducen, siendo una mejor opción si se les desea utilizar en cursos posteriores. Por último, respecto al tiempo de revisión, es claro que los exámenes computarizados y en línea son muy superiores a los tradicionales, ya que se califican de manera automática.

Finalmente, respecto al recurso de la ubicación, los exámenes en papel no dependen de la conexión de internet del centro escolar ni de los dispositivos usados por los estudiantes, siendo esta una de sus principales desventajas. Sin embargo, las pruebas en línea pueden ser realizados en cualquier lugar e incluso pueden resolverse vía celular, razón por la cual se les utilizó durante la contingencia para evaluar a los estudiantes a distancia.

Para terminar, una de las mayores debilidades de los exámenes en línea con respecto a los exámenes tradicionales se encuentra en el rubro de supervisión, ya que los docentes no pueden asegurarse de que los estudiantes no hayan hecho trampa comunicándose entre sí, consultando apuntes o buscando en internet. El nivel de esta cualidad con que cuentan los exámenes computarizados es bastante bajo cuando son realizados en casa, como sucedió durante el periodo estudiado, ya que exigen plena confianza en la honestidad de quien los contesta. Sin embargo, se puede afirmar que los exámenes computarizados, cuando son aplicados en un centro de cómputo al interior del centro escolar y vigilados por un docente, tendrían el mismo nivel de supervisión que los hechos en papel. Por otro lado, cabe mencionar la aportación de McVay, quien afirma que las estrategias usadas en la evaluación a distancia han de aplicarse bajo ciertos supuestos, entre los cuales están “renunciar al control, aplicación en el mundo real y uso de evidencias de aprendizaje” (2002, p. 123). En primera instancia, esa renuncia implicaría que el control de lo que se evalúa esté cada vez más en manos del estudiante, y no

del docente; mientras que los dos supuestos siguientes están directamente relacionados con el uso de elementos «suaves» de evaluación.

En general, los exámenes tradicionales obtuvieron un puntaje de 35.3 y los exámenes computarizados, de ser aplicados al interior de la institución, uno de 34.7. Si bien los exámenes tradicionales siguen mostrándose superiores a los computarizados, la diferencia no es tan amplia como la que existe si se les compara con los exámenes aplicados a distancia (31.7 puntos), donde el puntaje de supervisión juega un papel determinante.

## **V. CONCLUSIONES**

Para la realización de esta investigación, se comenzó con una revisión bibliográfica sobre la evaluación, la evaluación de los cursos de inglés, y los diferentes tipos de exámenes. A continuación, se procedió a encontrar una serie de cualidades que ayudaran a comparar las evaluaciones tradicionales con las computarizadas. Posteriormente, se analizó el proceso de evaluación que se llevaba a cabo previo al confinamiento y se le comparó con el realizado durante el semestre estudiado. Finalmente, se comparó a los exámenes tradicionales con los exámenes computarizados y los exámenes en línea, usando la lista de cualidades propuesta por Frodden et al., para así poder establecer qué tipo de exámenes presentan más beneficios.

Aunque se encontró que los exámenes tradicionales continúan considerándose superiores, también se demostró que existen cualidades en las que los exámenes computarizados son dominantes. De hecho, si se realizan bajo la misma supervisión docente que los exámenes tradicionales, la diferencia general se reduce de manera importante.

En ambos casos, la autenticidad y la interactividad obtuvieron puntajes bajos. Sin embargo, hay que recordar que, además de una prueba de tipo «duro», la evaluación de los cursos de inglés se complementa siempre con elementos de tipo «suave», que precisamente tienen mayor grado de ambas cualidades. Este es el caso de las habilidades productivas que, como ya se mencionó, fueron evaluadas a través de otro tipo de instrumentos.

A los profesores que opten por continuar utilizando pruebas tradicionales, se les propone que busquen mecanismos o estrategias para que estos sean más fiables y más prácticos. Por su parte, a aquellos docentes que deseen utilizar pruebas computarizadas, se les sugiere indagar sobre cómo elevar los niveles de impacto, supervisión y validez de la construcción, sobre todo en la habilidad de *writing*, así como buscar más opciones a través de las cuales puedan evaluarse las habilidades productivas.

## ANEXO 1.

### Evaluación mediante plataformas digitales

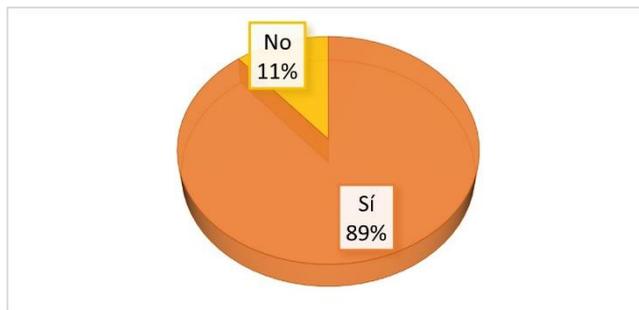
Creada con *Survey Monkey* el 15 de octubre de 2020.

Total de encuestados: 18

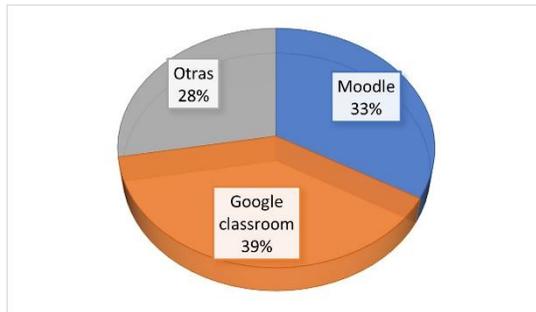
Vínculo con resultados (original en inglés):  
[https://es.surveymonkey.com/analyze/gQ\\_2FZtdfInW5uwsCxn7P44dWzeUf\\_k9MR4SIcYKxGMtEo\\_3D?tab\\_clicked=1](https://es.surveymonkey.com/analyze/gQ_2FZtdfInW5uwsCxn7P44dWzeUf_k9MR4SIcYKxGMtEo_3D?tab_clicked=1)

### Resultados:

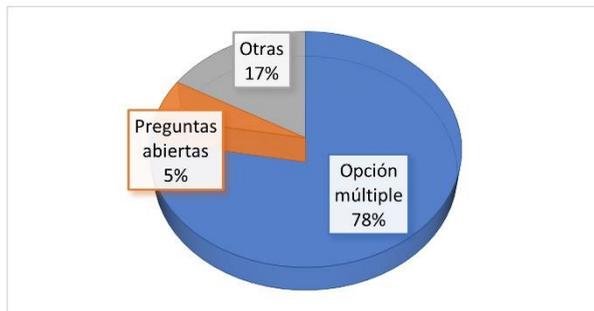
1. ¿Ha utilizado una plataforma virtual (como Moodle o Schoology) para evaluar a sus estudiantes?		
	Respuestas	Porcentaje
Sí	16	88.89%
No	2	11.11%



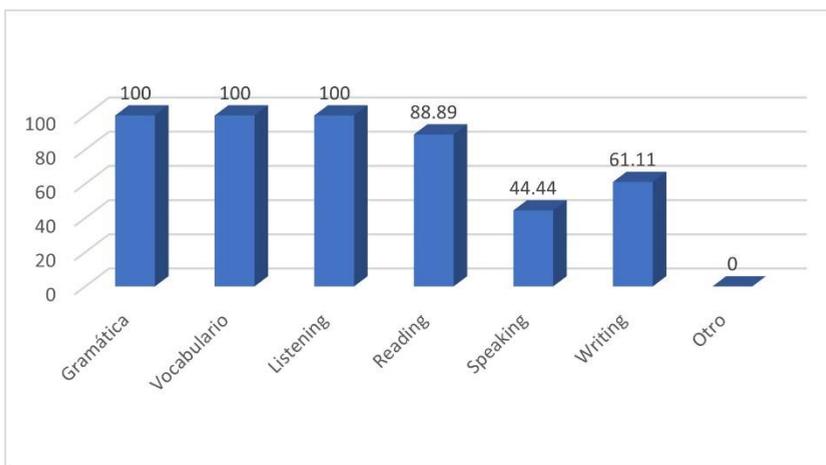
2. ¿Qué plataforma prefiere? (no se muestran aquellas que tienen 0 menciones)			
	Respuestas	Porcentaje	
Moodle	6	33.33%	Socrative (2), Edmodo, Canvas Instructure, Testmoz, Nearpod
Google Classroom	7	38.89%	
Otras	5	27.78%	



3. ¿Qué tipo de reactivos usa con mayor frecuencia?			
	Respuestas	Porcentaje	
Opción múltiple	14	78%	<i>Cloze, short answer, multiple choice grid, check grid, etc.</i>
Preguntas abiertas	1	5.56%	
Otras	3	16.67%	



4. ¿Qué elementos evalúa usando la plataforma virtual? (Puede marcar más de una opción)					
	Respuestas	Porcentaje		Respuestas	Porcentaje
Gramática	18	100	<i>Speaking</i>	8	44.44
Vocabulario	18	100	<i>Writing</i>	11	61.11
<i>Listening</i>	18	100	Otro	0	0
<i>Reading</i>	16	88.89			



5. ¿Cree que la evaluación a través de plataformas virtuales es confiable?		
	Respuestas	Porcentaje
Sí, totalmente	3	16.67%
Sí, pero no en un 100%	8	44.44%
Es más o menos confiable	6	33.33%
No es para nada confiable	1	5.56%



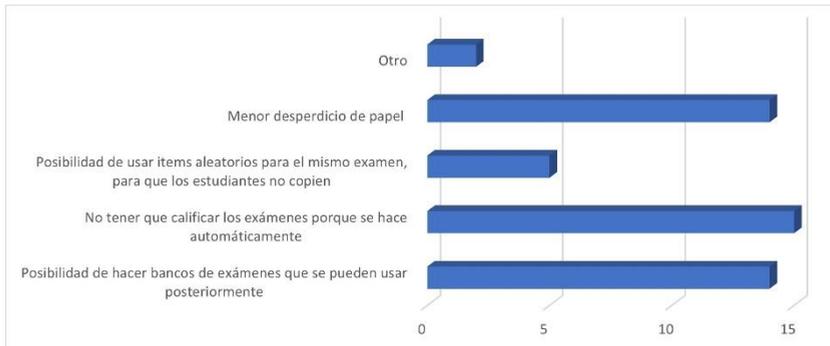
6. ¿Cuáles son las tres principales desventajas de usar plataformas virtuales para evaluar los aprendizajes?

	Respuestas	Porcentaje	
Falta de confiabilidad	9	52.94%	
Problemas de conexión de los estudiantes	14	82.35%	
Posibilidad de hacer trampa	14	82.35%	
Dificultad para controlar las veces que los estudiantes escuchan el audio	4	23.53%	
Falta de oportunidad para responder las preguntas de los estudiantes	6	35.39%	
Mi propia falta de habilidades en programación	2	11.76%	
Otro	1	5.88%	Depende de cómo sea usada la plataforma

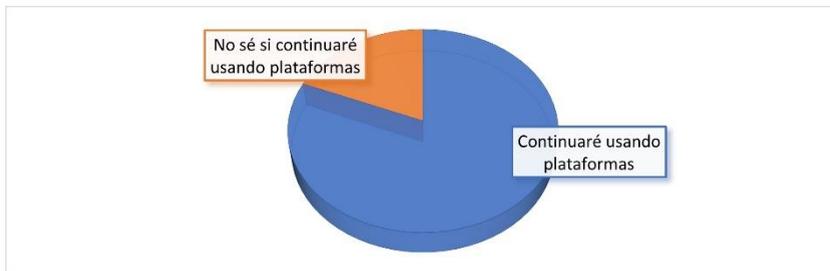


7. ¿Cuáles son las tres principales ventajas de usar una plataforma virtual para evaluar?

	Respuestas	Porcentaje	
Posibilidad de hacer bancos de exámenes que se pueden usar posteriormente	14	87.50%	
No tener que calificar los exámenes porque se hace automáticamente	15	93.75%	
Posibilidad de usar reactivos aleatorios para el mismo examen, para que los estudiantes no copien	5	31.35%	
Menor desperdicio de papel	14	87.50%	
Otro	2	12.50%	Permite a los estudiantes mostrar su trabajo, aunque no estén en la escuela; promueve el aprendizaje autónomo y la administración del tiempo, es flexible.



8. Cuando la pandemia termine y volvamos a las aulas...		
	Respuestas	Porcentaje
Definitivamente continuaré usando plataformas virtuales como parte de mi evaluación	13	81.25%
No sé si continuaré usando plataformas virtuales como parte de mi evaluación	3	18.75%
Dejaré de usar plataformas virtuales y volveré a mis herramientas de evaluación anteriores	0	0.00%



## REFERENCIAS

- Aoun, J. E. (2018). *Robot-Proof. Higher Education in the Age of Artificial Intelligence*. MIT Press.
- Aprender.org (2020). ¿Cuántos árboles se necesitan para hacer una hoja? [Mensaje en un blog] *Aprender.org*  
<https://www.aprender.org/cuantos-arboles-se-necesitan-para-hacer-una-hoja/>
- Castillo, S. & Cabrerizo, J. (2010). *Evaluación educativa de aprendizajes y competencias*. Pearson Prentice Hall.
- Consejo de Educación Inicial y Primaria [CEIP] (2020, marzo 23). Circular N°2. *Entremedios: construir cercanía en la distancia*.  
[http://www.ceip.edu.uy/documentos/normativa/tecnica/2020/Circular2\\_20\\_TECNICA.pdf](http://www.ceip.edu.uy/documentos/normativa/tecnica/2020/Circular2_20_TECNICA.pdf)
- Frodden, M.C.; Restrepo, M. I. & Maturana, L. (2004). Analysis of Assessment Instruments Used in Foreign Language Teaching. *Íkala*, 9(15), pp. 171-201. <https://revistas.udea.edu.co/index.php/ikala/article/view/3146>
- Harmer, J. (2014). *How to teach English*. Pearson.
- Hernández, R.; Fernández, C.; Baptista, M. P.; Méndez, S. & Mendoza, C. P. (2014). *Metodología de la investigación*. McGrawHill.
- McVay, M. (2002). *The Online Educator: A Guide to Creating the Virtual Classroom*. Routledge/Falmer.
- Morgan, C. & O'Reilly, M. (2002). *Assessing Open and Distance Learners*. Kogan Page.
- Quesada, R. (2006). Evaluación del aprendizaje en la educación a distancia “en línea”. *RED. Revista de Educación a Distancia*, (Número especial dedicado a la evaluación en entornos virtuales de aprendizaje), pp. 1-15. <https://www.um.es/ead/red/M6/quesada.pdf>

Leonora Estrada Flores / Mabel García Mendoza / Esmeralda Irais León López /  
Mónica Loya Rangel / María Sánchez Santiago / Luis Fernando González  
Valoración de la evaluación del curso de inglés  
a través de exámenes computarizados  
Revista *Xihmai* XVII (33), 9-42, enero-junio 2022

- Rodríguez, N. (2014). Fundamentos del proceso educativo a distancia: enseñanza, aprendizaje y evaluación. *RIED. Revista Iberoamericana de Educación a Distancia*, 17(2), pp. 75-93.  
<https://doi.org/10.5944/ried.17.2.12679>
- Ryan, S.; Scout, B.; Freeman, H. & Patel, D. (2000). *The Virtual University*. Kogan Page.
- Uribe, C.; Brito, L. & Müller-Marín, K. (2020, marzo 25). En tiempos de crisis, la UNESCO está presente. [Mensaje en un blog] *UNESCO*.  
<https://es.unesco.org/fieldoffice/santiago/articles/la-unesco-esta-presente>

Copyright (c) 2022 Leonora Estrada Flores, Mabel García Mendoza, Esmeralda Irais León López, Mónica Loya Rangel, María Sánchez Santiago, Luis Fernando González Juárez.



Este texto está protegido por una licencia [Creative Commons 4.0](https://creativecommons.org/licenses/by/4.0/)

Usted es libre de:

- 1) Compartir — copiar y redistribuir el material en cualquier medio o formato.
- 2) Adaptar — remezclar, transformar y construir a partir del material para cualquier propósito, incluso comercialmente, siempre que cumpla la condición de: **Atribución** — Usted debe dar [crédito de manera adecuada](#), brindar un enlace a la licencia, e [indicar si se han realizado cambios](#). Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que usted o su uso tienen el apoyo de la licenciante.

[ResumenDeLicencia](#)

[TextoCompletoDeLicencia](#)

Leonora Estrada Flores / Mabel García Mendoza / Esmeralda Irais León López /  
Mónica Loya Rangel / María Sánchez Santiago / Luis Fernando González  
Valoración de la evaluación del curso de inglés  
a través de exámenes computarizados  
Revista *Xihmai* XVII (33), 9-42, enero-junio 2022