



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

INSTRUMENTO DE EVALUACIÓN DINÁMICA BASADO EN CÓMICS PARA MEDIR LA COMPRENSIÓN LECTORA EN ALUMNOS DE TERCERO BÁSICO

IGNACIO JOSÉ CARMACH ITURRIETA

Tesis para optar al grado de
Magíster en Ciencias de la Ingeniería

Profesor Supervisor:
MIGUEL NUSSBAUM VOEHL

Santiago de Chile, (Junio, 2016)

© 2016, Ignacio Carmach Iturrieta



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

INSTRUMENTO DE EVALUACIÓN DINÁMICA BASADO EN CÓMICS PARA MEDIR LA COMPRENSIÓN LECTORA EN ALUMNOS DE TERCERO BÁSICO

IGNACIO JOSÉ CARMACH ITURRIETA

Tesis presentada a la Comisión integrada por los profesores:

MIGUEL NUSSBAUM VOEHL

LORENA MEDINA MORALES

YADRAN ETEROVIC SOLANO

VLADIMIR MARIANOV KLUGE

Para completar las exigencias del grado de
Magíster en Ciencias de la Ingeniería

Santiago de Chile, (Junio, 2016)

A mis padres, mi familia, Valentina y
mis amigos, por su apoyo
incondicional.

AGRADECIMIENTOS

A mis padres, por su constante apoyo durante todas las etapas de mi vida y cuyo esfuerzo me ha permitido tanto.

A mis hermanos, que han estado siempre presentes en mi vida y me apoyaron durante todos los procesos de este trabajo.

A Valentina Pérez, por hacer de mí una mejor persona cada día y por su incondicional apoyo y preocupación, sin los cuales este proyecto no se habría completado.

A Juan Jaime Díaz, Tomás Martínez y Enzo Sciarresi, por los increíbles momentos que hemos vivido juntos y todos los que vendrán, por su constante apoyo y motivación a seguir adelante. A Juan Jaime por su gran aporte al desarrollo de esta investigación.

Al profesor Miguel Nussbaum, por su constante preocupación y energía, por acercarnos al mundo de la investigación desde tan temprano y motivarnos a no quedarnos quietos frente a los problemas.

INDICE GENERAL

	Pág.
DEDICATORIA	ii
AGRADECIMIENTOS	iii
INDICE DE TABLAS	vi
INDICE DE FIGURAS.....	vii
RESUMEN.....	viii
ABSTRACT	ix
1. Introducción.....	1
1.1 Objetivos e hipótesis	6
2. Metodología.....	8
3. Instrumento de Evaluación Dinámica.....	10
4. Arquitectura de software.....	21
4.1 Capa de presentación.....	22
4.2 Capa de negocio	24
4.3 Capa de datos	26
5. Experimentación en colegio	30
5.1 Primera iteración	31
5.2 Segunda iteración	32
6. Resultados.....	35
6.1 Primera iteración	35
6.2 Segunda iteración	36
7. Conclusiones.....	39
BIBLIOGRAFIA	41

A N E X O S	44
Anexo A: Artículo publicado por la revista Interactive Learning Environments	45
Anexo B: Artículo enviado a la revista Journal of Research in Reading.....	85
Anexo C: Preguntas diseñadas para el instrumento	118
Anexo D: Resultados de los distintos instrumentos utilizados en la investigación	123

INDICE DE TABLAS

	Pág.
Tabla 1-1: Características del instrumento relacionadas con los objetivos.....	7
Tabla 3-1: Extracto de las preguntas de cada sección.....	12
Tabla 5-1: Notas y tiempos obtenidos por los alumnos en el instrumento	34
Tabla 6-1: Análisis del índice de confiabilidad por sección	36
Tabla 6-2: Regresiones lineales en función de los resultados del instrumento	37
Tabla 6-3: Test ANOVA de dos factores para determinar sesgo de género	38

INDICE DE FIGURAS

	Pág.
Figura 2-1: ILD Framework. (Roschelle, 2010; Bannan-Ritland, 2003).....	9
Figura 3-1: Vista de la pantalla de preguntas al final de las secciones	13
Figura 3-2: Flujo entre niveles de dificultad cuando no existe error acumulado	15
Figura 3-3: Flujo entre niveles de dificultad con error acumulado igual a 1	16
Figura 3-4: Vista de la pantalla general del instrumento	17
Figura 3-5: Vista de la pantalla de inicio de las secciones.....	18
Figura 3-6: Ejemplo de reporte de desempeño del instrumento	19
Figura 4-1: Arquitectura de software del instrumento	21
Figura 4-2: Estructura básica del código HTML del instrumento	22
Figura 4-3: Diagrama de secuencia del funcionamiento de la capa de negocio	25
Figura 4-4: Diagrama de la base de datos del instrumento	29
Figura 5-1: Distribución y montaje del instrumento durante la experimentación	32
Figura 5-2: Aplicación del instrumento durante la experimentación.....	33

RESUMEN

El uso de la tecnología en la evaluación de la comprensión lectora no ha sido extensivo en el último tiempo, limitándose a la digitalización de pruebas estandarizadas. Su evaluación requiere múltiples indicadores para medir las variables subyacentes y evitar problemas de validez de constructo. La evaluación dinámica puede ser una alternativa interesante a las pruebas tradicionales, pues es capaz de medir y adaptarse al nivel del estudiante. Por otro lado, un alumno se acerca a un cómic por una motivación intrínseca, lo que lo hace especialmente atractivo como instrumento de medición.

Para esta investigación se desarrolló un instrumento de evaluación dinámica, basado en cómics, para medir la comprensión lectora en alumnos de tercero básico. El instrumento se diseñó para ser usado en tablets y luego se validó internamente, y externamente con una prueba estandarizada y las notas de los alumnos en Lenguaje y Comunicación.

La experimentación realizada se llevó a cabo utilizando el software mencionado en cursos donde previamente se había aplicado la prueba SEPA, desarrollada por MIDE-UC, de forma independiente por los profesores de cada establecimiento.

Los resultados del instrumento son comparables con aquellos que pueden entregar los métodos de evaluación tradicionales, por medio de pruebas estandarizadas. Además, el software es capaz de entregar estos resultados en tiempo real y hacerse accesible a un gran número de estudiantes a un muy bajo costo a través de internet.

Esta tesis cuenta con el apoyo del proyecto FONDECYT 1150045.

Palabras Claves: comprensión lectora, evaluación dinámica, cómic, aprendizaje interactivo

ABSTRACT

The use of technology in the assessment of reading comprehension has not been extensive in recent times, limited to the digitalization of standardized tests. Its evaluation requires multiple indicators in order to measure underlying variables and avoid issues with construct validity. Dynamic assessment methods may provide a good alternative to traditional tests, as they are capable of adapting to and measuring a student's level. In addition to this, students are drawn to comics through intrinsic motivation, making them attractive as an assessment instrument.

For this research a comic-based dynamic assessment instrument was developed for measuring reading comprehension among third grade students. The instrument was designed for use with tablets and then validated internally, and externally using a standardized test and the students' grades in Language Arts.

The experiment was carried out using the mentioned software in courses where the SEPA test, developed by MIDE-UC, was previously and independently applied by the teachers of each establishment.

After the experimentation, it can be observed that the results of the instrument are comparable with those delivered by traditional assessment methods, like standardized tests. In addition, the software is able to deliver those results in real time and become accessible to a large number of students at a very low cost through internet.

This thesis is supported by FONDECYT project 1150045.

Keywords: reading comprehension, dynamic assessment, comic, interactive learning

1. INTRODUCCIÓN

Actualmente, la tecnología ha irrumpido en casi todos los aspectos de nuestras vidas, todo lo que hacemos o usamos está de alguna forma relacionado a ella. Los libros y la lectura no han quedado ajenos a esto y, hoy en día, el acceso universal a vastas bibliotecas está dado en gran medida gracias al uso de la misma. A pesar de lo anterior y considerando que Chile tiene el mejor puntaje de América Latina y el Caribe en lectura y matemáticas, para tercero básico (TERCE, Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, 2015), un estudio revela que el 49% de los chilenos dice no leer libros (CERLALC - UNESCO, 2012). Si a esto sumamos que el 44% de su población no comprende lo que lee (Centro de Microdatos, Departamento de Economía, Universidad de Chile, 2013) y que el 33% de sus estudiantes no alcanza las competencias mínimas en lectura para desenvolverse en el mundo actual y futuro (OCDE 2014), vemos que la situación es claramente deficiente.

Al considerar la realidad de otros países, encontramos, por ejemplo, que en Estados Unidos más de una tercera parte de la población tiene problemas de lectura, a tal grado que se estima que 36 millones de adultos norteamericanos son analfabetos funcionales (Farrell, 2015). Así pues, notamos que aún en los países clasificados como de alto desarrollo, los problemas de lectura están presentes en el contexto de sus sociedades, haciendo patente la necesidad de herramientas que ayuden a evaluarlos y corregirlos.

Por esto nace en mí la inquietud de buscar soluciones tecnológicas que se enfoquen en mejorar la calidad y aprendizaje de los contenidos, más que en la disponibilidad de los

mismos. Es así como en el año 2009, durante mi tercer semestre en la universidad, conozco al profesor Miguel Nussbaum, quien se encontraba dirigiendo el proyecto de doctorado de Juan Felipe Calderón. Este trabajo buscaba mejorar el laboratorio tradicional de idiomas al basarlo en Single Display Groupware (SDG) para el aprendizaje de vocabulario, gramática, pronunciación y comprensión auditiva (Calderón et al., 2014). Al adoptar SDG, el laboratorio de idiomas se beneficia de las ventajas del aprendizaje colaborativo en grupos pequeños. Dada esta inquietud que tenía, cuando me ofrecen participar de este proyecto, no lo dudé y me uní al equipo de trabajo. Durante los siguientes 3 años de trabajo colaboré en esta investigación, comenzando por desarrollar software que permitiera enviar distintos audios a más de una tarjeta de sonido de forma simultánea, facilitando así el trabajo colaborativo de los grupos. Luego pasé al desarrollo de las interfaces de usuario de los módulos de vocabulario, gramática, pronunciación y comprensión auditiva del laboratorio, teniendo especial cuidado en la usabilidad de las mismas, ya que el software estaba destinado a alumnos del primer ciclo básico. Una vez que el desarrollo del software finalizó pasé a formar parte del equipo que visitó los colegios para hacer las pruebas en terreno. Como apéndice al presente documento (Anexo A) se encuentra un paper publicado por la revista *Interactive Learning Environments*, en el que se exponen las principales conclusiones obtenidas a lo largo de aquella investigación.

A nivel personal, esta colaboración sienta las bases de mi trabajo futuro y me impulsa a buscar nuevas soluciones enfocadas en la educación, además de proporcionar las herramientas y experiencias necesarias para llevar a cabo una investigación propia, tanto a nivel teórico como a nivel práctico. Es así como, junto a Juan Jaime Díaz, proponemos

al profesor Nussbaum explorar alternativas en torno a la evaluación del lenguaje, pero esta vez sin enfocarnos en un idioma en particular.

Cuando pensamos en lenguaje, la evaluación y desarrollo de la comprensión lectora surge como uno de los temas principales tanto a nivel nacional como internacional. Evidencia de esto último es que existen iniciativas gubernamentales que año a año miden el desempeño de los alumnos de cada colegio, por ejemplo, en Chile, la prueba SIMCE¹ se realiza todos los años a estudiantes de 4° y 8° Básico. Sin embargo, los resultados de esta evaluación se reciben cerca de seis meses después de realizada, lo que dificulta la corrección de las deficiencias identificadas. Otras, como el Informe PISA² (OCDE, 2011) tienen períodos aún más largos, aplicándose cada 3 años y, además, examina a estudiantes de una determinada edad y no de un nivel específico. También existen iniciativas privadas, por ejemplo la prueba SEPA³ (MIDE-UC, 2009a), que se realiza por encargo de los colegios hasta 2 veces al año a los alumnos de 1° Básico a III Medio. En este caso la entrega de resultados puede demorar meses.

Al considerar específicamente la comprensión lectora, debemos entender que es un constructo multifacético. La elección de los textos y tareas que se incluyen en su evaluación conducirán, en cierta medida, los resultados que se obtengan (O'Reilly et al., 2014). La evaluación requiere incorporar múltiples indicadores que permitan precisión en la medición de las variables subyacentes (Fletcher, 2006). Si esto no ocurre, los resultados de la evaluación pueden tener sesgo o invalidez del constructo (Cook y

¹ SIMCE: Sistema de Medición de la Calidad de la Educación. Más información en: <http://www.simce.cl>

² PISA: Programme for International Student Assessment. Más información en: <http://www.oecd.org/pisa/aboutpisa/>

³ SEPA: Sistema de Evaluación de Progreso del Aprendizaje. Más información en: <http://www.sepauc.cl>

Campbell, 1979). Los métodos de evaluación de la comprensión lectora no han sufrido grandes cambios a través de los años. Se realizan principalmente por medio de pruebas estandarizadas de papel y lápiz. Dada la naturaleza estática de las mismas, sólo son capaces de identificar las habilidades ya existentes de los estudiantes y no consideran las diferencias entre cada uno de ellos (Nazari y Mansouri, 2014). Esto último es relevante si se considera que, hoy en día, la literatura internacional evidencia que las prácticas efectivas de lectura inicial incorporan las diferencias individuales de los estudiantes (Connor, Morrison y Katch, 2004; Snow, Griffin y Burns, 2005; Connor et al., 2009).

En contraste, los métodos de evaluación dinámicos son capaces de adaptarse a cada estudiante y a su nivel de conocimiento, pudiendo incluso determinar el potencial de aprendizaje de estos, lo que Vygotsky llama zona de desarrollo próximo (ZDP) (Vygotsky, 1978). Las mediciones estáticas sólo muestran lo ya aprendido o logrado, mientras que la amplitud de la ZDP proporciona posibles indicios de lo que se puede aprender. Así las mediciones dinámicas dan la posibilidad de medir y entender el nivel del estudiante, además de adaptarse y mejorar desde ahí su nivel de desarrollo. Para una comparación entre los métodos estáticos y dinámicos se sugiere ver Tabla 1 en (Nazari y Mansouri, 2014).

El uso de la tecnología en la evaluación de la comprensión lectora no ha sido extensivo en el último tiempo, en la mayoría de los casos limitándose a llevar las pruebas estandarizadas del papel al mundo digital. A pesar de esto, surge la inquietud por nuevas alfabetizaciones, a partir de las oportunidades que otorgan las nuevas tecnologías, soportes que sirven al despliegue de textos híbridos o discontinuos y multimodales

(Cope y Kalantzis, 2009). El resultado de esta digitalización, ha sido ir valorando e incorporando, más allá del escrito impreso en papel, lo visual y otros modos expresivos. La combinación de la inquietud por incorporar tecnologías de forma poco convencional al estudio del lenguaje con nuestros gustos personales por las novelas gráficas o cómics, nos lleva a considerar estos últimos como un tipo de texto viable para medir la comprensión lectora. El cómic resulta ser un recurso adecuado dada la reacción de los estudiantes frente a este. La incorporación de gráficos visuales e imágenes junto con el texto pueden servir no solo para motivar a los estudiantes, sino también para el desarrollo de la alfabetización visual (Valerie y Abed, 2013). Un alumno se acerca a un cómic por una motivación intrínseca, especialmente si se adapta a sus intereses (Cimermanová, 2015). El cómic cumple con las siguientes características que lo hacen especialmente válido como instrumento de medición (Cuadrado et al., 1999): es más fácil de comprender que los textos verbales, gusta tanto a profesores como a alumnos, es divertido, provoca, atrae, impacta, fomenta la creatividad, favorece la dinámica de clase, llega en un golpe de vista, y dan color a la clase. Además de lo anterior, la multimodalidad del cómic se puede implementar bien al medio digital y éste puede estimular el interés de los estudiantes por la lectura (Frey y Fisher, 2004).

A pesar de lo anterior, aún no se han evidenciado experiencias basadas en el uso de cómics en formato digital que busquen la evaluación de la comprensión lectora, por lo que a lo largo de este documento de tesis se analizará el desempeño de este tipo de instrumento.

Las siguientes secciones presentan el detalle de la investigación realizada, en la que se diseñó y desarrolló un software de evaluación dinámica de la comprensión lectora,

basado en el uso de cómics en formato digital. Se presentan además las características principales del sistema que son relevantes para la investigación, como las reglas de flujos entre niveles de dificultad, así como también el contraste de los resultados del instrumento frente a los de otras pruebas estandarizadas. Finalmente se realiza una explicación detallada de la experiencia y un análisis de los resultados obtenidos. Como apéndice al presente documento (Anexo B) se encuentra un paper enviado a la revista *Journal of Research in Reading*, en el que se exponen las principales conclusiones obtenidas a lo largo de esta investigación.

1.1 Objetivos e hipótesis

A raíz de lo expuesto anteriormente, se espera entonces realizar una experiencia que permita determinar si existe una correlación estadística entre los resultados entregados de forma inmediata por un instrumento especialmente diseñado para esta investigación y los obtenidos por alumnos de tercero básico en una prueba estandarizada de evaluación de la comprensión lectora.

La hipótesis de esta investigación es que los resultados obtenidos a través de un software de evaluación dinámica de la comprensión lectora, basado en el uso de cómics en formato digital, son comparables con aquellos que pueden entregar los métodos de evaluación tradicionales, por medio de pruebas estandarizadas. Además, se pretende evaluar si el set de reglas utilizado para determinar el flujo de los estudiantes a través de distintos niveles de dificultad en un software educativo, diseñado especialmente para la investigación, determina correctamente el nivel de comprensión lectora en el que se sitúa cada alumno.

En la Tabla 1-1 se observan las características del software desarrollado que están directamente relacionadas con los objetivos de la investigación.

Tabla 1-1: Características del instrumento relacionadas con los objetivos

Objetivo	Aspecto del instrumento
Correlación estadística entre los resultados del instrumento y las notas obtenidas por los alumnos en una prueba estandarizada.	Reportes del instrumento de evaluación dinámica.
Determinación del nivel de comprensión lectora en el que se encuentra cada alumno.	Reglas de flujo entre niveles de dificultad del instrumento.

2. METODOLOGÍA

Para el desarrollo de esta investigación se utilizó una metodología denominada Design-Based Research (DBR), cuyo objetivo es estudiar las posibilidades de nuevos entornos de aprendizaje y enseñanza en un ambiente real (The Design-Based Research Collective, 2003). A través de esta metodología y durante todo el proceso de diseño de un nuevo instrumento, los investigadores son capaces de conocer el ambiente donde finalmente este será utilizado. Además permite realizar estudios donde se analizan diversos factores de forma rápida y sistemática, creando procesos iterativos de mejora (Roschelle, 2010). En particular, esta investigación estuvo basada en el Diseño de Aprendizaje Integral (ILD, por sus siglas en inglés), marco flexible y amplio para la investigación con raíz en los principios de Design-Based Research (Bannan-Ritland, 2003). Como se observa en la Figura 2-1 (Roschelle, 2010; Bannan-Ritland, 2003), esta metodología consta de 4 etapas iterativas:

- 1) Exploración informada: Etapa donde se realiza un estudio de las necesidades, teorías y audiencia para definir el problema.
- 2) Diseño e implementación: Se realiza el desarrollo de un prototipo.
- 3) Evaluación local: Evaluación del impacto en los usuarios de la intervención realizada. En caso de requerir cambios se debe volver a la etapa anterior.
- 4) Evaluación final: Publicación, difusión y adopción de la nueva herramienta.

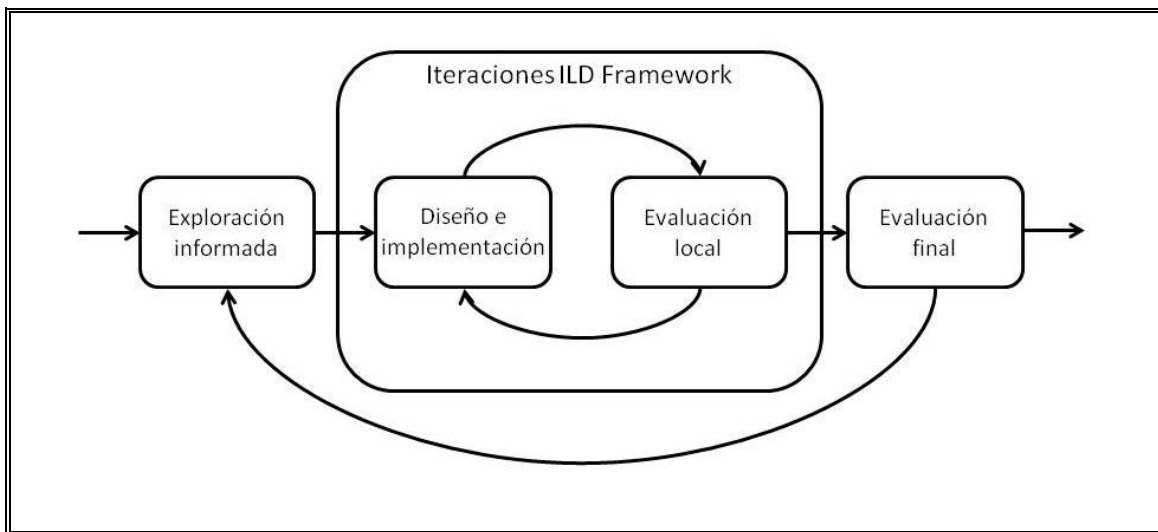


Figura 2-1: ILD Framework. (Roschelle, 2010; Bannan-Ritland, 2003)

Siguiendo estas etapas, en esta investigación primero se definió el problema y se plantearon los objetivos, mencionados en el punto anterior. Para cumplir con estos objetivos se desarrolló e implementó un instrumento de evaluación dinámica, basado en cómics, para medir la comprensión lectora en alumnos de tercero básico. El instrumento se diseñó para ser usado en tablets a través de una aplicación web. Posterior al desarrollo del mismo se realizaron dos iteraciones, en la primera se analizó el comportamiento de la herramienta con usuarios reales, en este caso alumnos de tercero básico de un colegio de Santiago, experiencia en la cual se logró identificar algunos problemas de usabilidad y se verificó la dificultad de las preguntas presentadas a lo largo del instrumento. Con estos resultados se comenzó la segunda iteración, donde se rediseñaron las secciones del instrumento que presentaban problemas y se corrigieron los errores observados, para luego volver a probar con usuarios reales. Esta vez durante la prueba se verificaría el cumplimiento de los objetivos principales y la eficiencia del nuevo diseño.

3. INSTRUMENTO DE EVALUACIÓN DINÁMICA

Como se mencionó en la sección anterior, durante la investigación se diseñó y desarrolló un instrumento de evaluación dinámica para medir la comprensión lectora en alumnos de tercero básico. Este instrumento se basa en un cómic, cuya narrativa multimodal corresponde a la adaptación de uno de los textos recomendados por el Ministerio de Educación de Chile para los estudiantes de 1° a 6° Básico: “La Vuelta al Mundo en 80 Días” de Julio Verne (MINEDUC, 2012), “La Vuelta al Mundo en 80 Días” consta originalmente de 42 capítulos, que fueron condensados en 17 secciones, desarrolladas en viñetas de cómic, que siguen la misma secuencia natural de la historia. El cómic como género multimodal, presenta pantallas de formato fijo, preferentemente horizontal, aunque también pueden aparecer en posición vertical; los personajes pueden ser fijos o no, quienes predominantemente utilizan diálogos en una narrativa secuenciada (Evans, 2013).

De acuerdo con la investigación contemporánea, las habilidades que se requieren para comprender un texto-discurso son multi-componenciales, operando en niveles de procesamiento complementarios: sintáctico, léxico-semántico y discursivo, como el tipo de texto y/o el género discursivo (Koda, 2005). Por ello, para el diseño y elaboración del cómic, se crearon matrices de reglas propias de las historias (categorías narrativas) y matrices de carácter textual, léxico-gramatical. Lo que se cruzó con tres niveles de complejidad lectora: básico, medio y avanzado, para la posterior construcción de las preguntas para cada nivel.

En detalle, ello significó, en primer término, la elaboración de una matriz de reglas del nivel macro, al modo de una “gramática de las historias”, para lo cual se consignaron categorías narrativas básicas, tales como escenario, personajes, tema y trama. Posteriormente, la historia se escribió en los tres niveles de dificultad señalados, a partir de una matriz de reglas, del nivel macro (oracional/enunciado) de complejidad gramatical y léxica, que permitió la producción de tres historias paralelas, con diferente extensión y complejidad. El léxico utilizado fue tomado de los listados sugeridos por el Ministerio para el nivel escolar seleccionado. Además, estas matrices sirvieron para la elaboración de una matriz de evaluación que permitiera construir las preguntas para cada nivel de la historia, y poder controlar así que la dificultad sólo estuviera dada por la comprensión del texto y no por la complejidad sintáctica y/o léxica de la formulación de las preguntas. Cada nivel tiene su propio tipo de preguntas asociadas: para el nivel básico, preguntas literales o textuales; para el nivel intermedio de inferencia simple, y para el nivel avanzado, de inferencia global o compleja, respectivamente. Las preguntas literales o textuales, suponen localizar información explícita en el texto, en la imagen, o en la relación texto – imagen, la que además puede aparecer incluso destacada de alguna manera en el texto. Las preguntas de inferencia simple, suponen respuestas que demandan la comprensión del significado de enunciados y palabras utilizando el contexto, estableciendo relaciones de causa-efecto, realizando comparaciones y contrastes, entre otros procesos. Las preguntas de inferencia compleja a menudo se centran en inferir a partir de pistas textuales, las ideas principales, secundarias, predecir acciones, conclusiones o resultados (León y Pérez, 2003). Estos niveles se corresponden con los que se evalúan a nivel macro en la mayor parte de las pruebas estandarizadas y

validadas a nivel internacional, tales como PISA y PIRLS, y la prueba SEPA (MIDE-UC, 2009b, 2009c), utilizada en esta investigación.

Tabla 3-1: Extracto de las preguntas de cada sección

ID	Sección	Nivel de dificultad	Pregunta
1	1	Básico	¿En qué ciudad comienza la historia?
7	2	Intermedio	¿Quién es el nuevo jefe de Picaporte?
14	3	Avanzado	¿Qué profesión tiene el señor Fix?
16	4	Básico	¿Por qué Fogg se va a dormir?
23	5	Intermedio	¿Cómo deciden viajar para llegar más rápido a Calcuta?
30	6	Avanzado	¿Por qué razón detienen a Fogg y a Picaporte?
31	7	Básico	¿Quién descubre que el barco está en reparaciones?
37	8	Intermedio	¿Por qué Picaporte se queda dormido en el bar?
44	9	Avanzado	¿Por qué Fogg perdió el barco a Yokohama?
46	10	Básico	¿Para qué se une Picaporte al circo?
53	11	Intermedio	¿Dónde se reencuentran Fogg y Picaporte?
60	12	Avanzado	¿Por qué Picaporte golpea a Fix?
61	13	Básico	¿Cuánto se retrasó el tren por causa de los bisontes?
67	14	Intermedio	¿Por qué Fogg sale a buscar a Picaporte?
74	15	Avanzado	¿A quién se le ocurre usar madera del mismo barco para la caldera?
76	16	Básico	¿Dónde arresta Fix a Fogg?
83	17	Intermedio	¿Por qué Fogg cree que perdió la apuesta?
85	17	Avanzado	¿Por qué Fogg y Picaporte ganaron un día en su viaje?

En relación a lo anterior y de acuerdo con el nivel en el cual se encuentra el alumno, las preguntas que se le hacen a lo largo del instrumento van variando de dificultad. De un total de 85 preguntas diseñadas (Anexo C), en la Tabla 3-1 se muestra un extracto para cada sección.

Una vez que el alumno termina una sección, un personaje le realiza una pregunta y le presenta tres alternativas, de las cuales debe seleccionar una. En la Figura 3-1 se muestra un ejemplo de un cuadro donde se realizan preguntas.

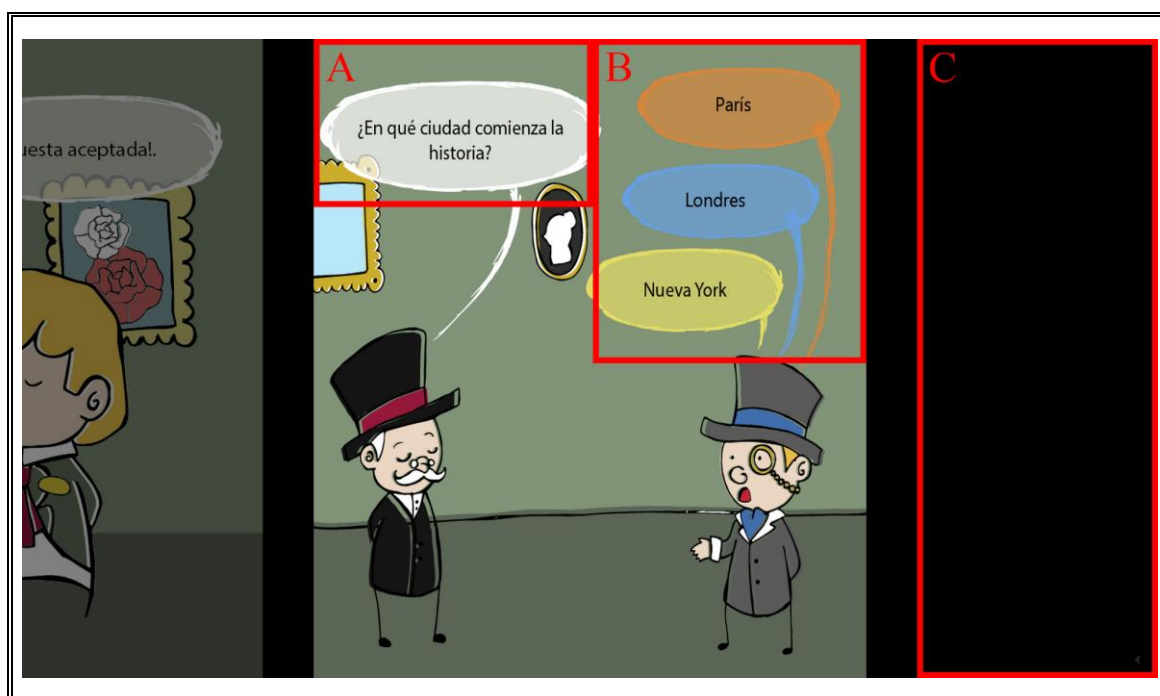


Figura 3-1: Vista de la pantalla de preguntas al final de las secciones

Cuando el alumno contesta correctamente una pregunta progresa a la siguiente sección, en la que los textos y las preguntas asociadas tienen una dificultad mayor, o se mantendrá la dificultad si el alumno se encuentra en el nivel avanzado. Cuando el

estudiante comete un error debe volver a leer la misma sección, pero esta vez los textos y la pregunta tienen una dificultad menor. En caso que cometa dos errores consecutivos, el sistema disminuye nuevamente el nivel de dificultad, o se mantiene en el nivel básico si ya se encontraba en éste. El flujo entre los distintos niveles de dificultad se detalla en las Figuras 3-2 y 3-3.

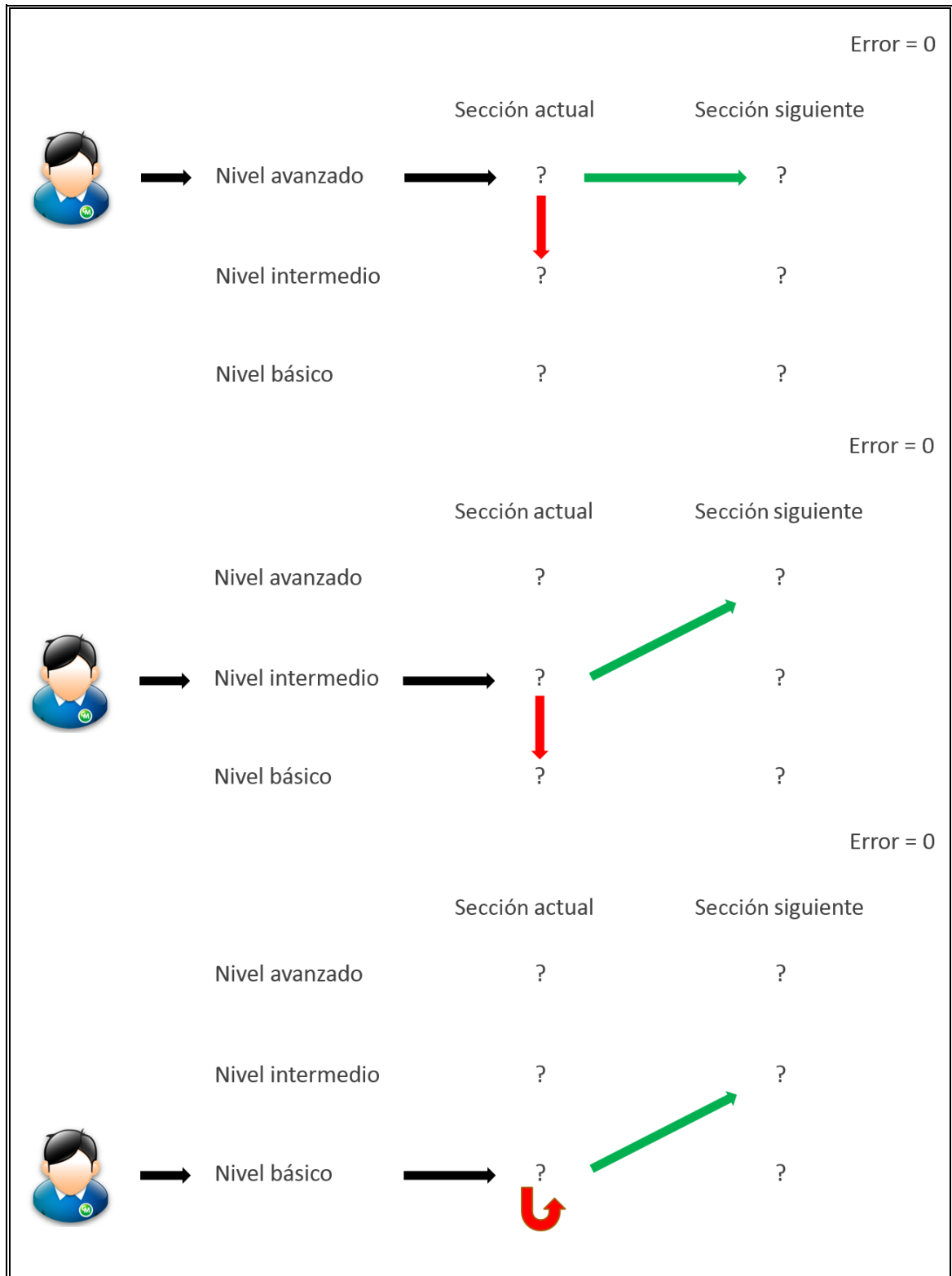


Figura 3-2: Flujo entre niveles de dificultad cuando no existe error acumulado

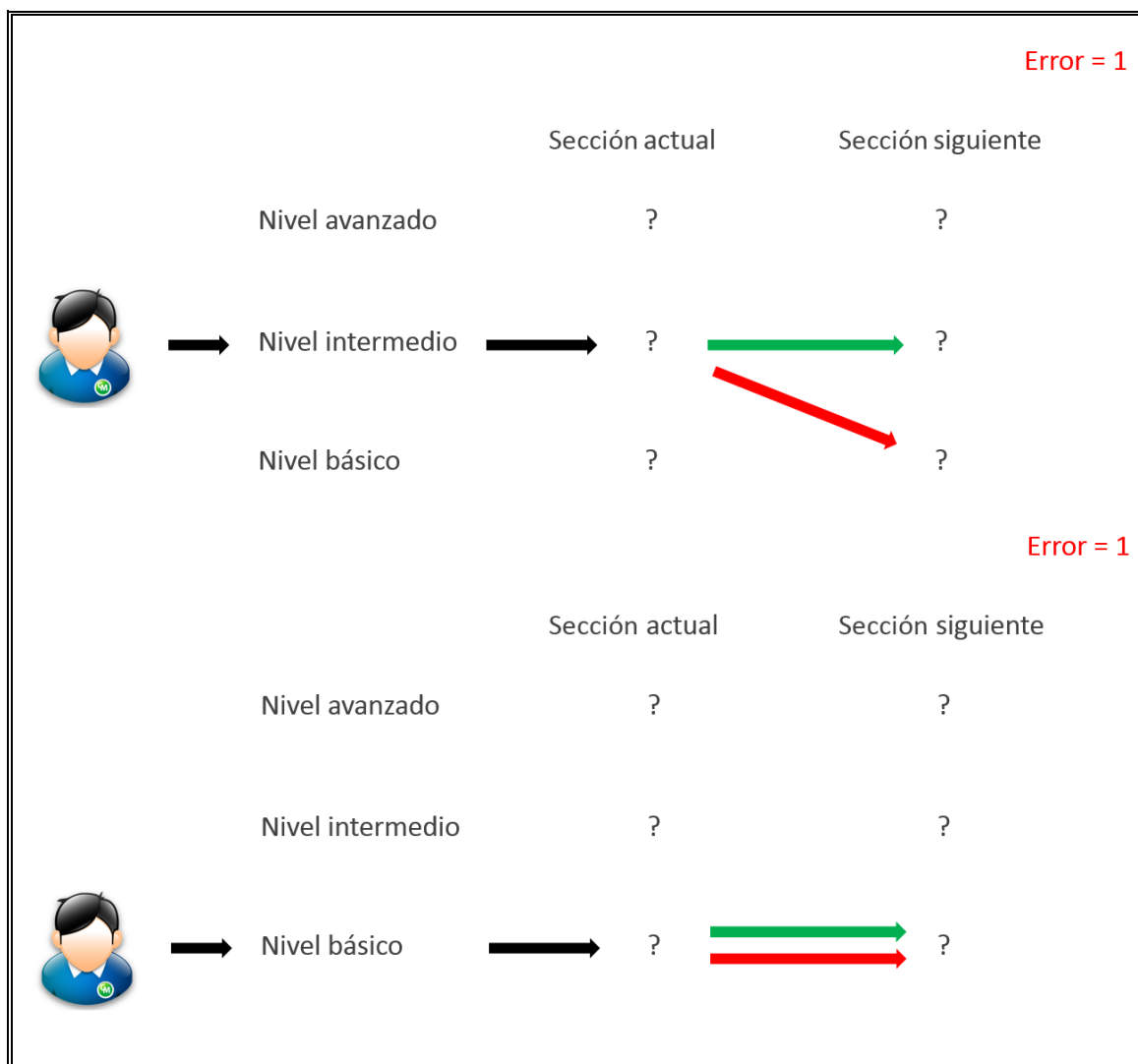


Figura 3-3: Flujo entre niveles de dificultad con error acumulado igual a 1

La extensión de la historia que finalmente es leída por un estudiante depende directamente de su nivel de comprensión lectora y de la cantidad de errores que cometa durante la actividad.

La sesión de evaluación comienza con la identificación del estudiante, quien elige su nombre de una lista previamente configurada. El sistema va guardando su progreso a medida que va avanzando en las secciones del cómic. Esto permite interrumpir la lectura

y salir del sistema y luego retomar en la misma sección donde estaba, además de poder avanzar el alumno a su propio ritmo y de acuerdo con sus propias necesidades.



Figura 3-4: Vista de la pantalla general del instrumento

Como se ve en la Figura 3-4, la visualización del instrumento se organiza en 3 áreas de la pantalla:

- 1) Viñeta anterior (Figura 3-4, área “A”): Espacio que muestra parcialmente el contenido de la última viñeta que se acaba de leer. La intención es dar un aspecto visual cercano a una tira cómica. Al comienzo de cada sección esta área está vacía (Figura 3-5, área “A”).
- 2) Contenido principal (Figura 3-4, área “B”): Espacio donde se muestra el contenido que actualmente está leyendo el alumno, que corresponde a la viñeta

donde se encuentra. Cuando el alumno llega al final de cada sección, esta área cambia con el fin de realizar la pregunta de comprensión lectora. En la parte superior izquierda del área (Figura 3-1, área “A”) se muestra la pregunta y en la parte superior derecha (Figura 3-1, área “B”), las 3 opciones de respuesta.

- 3) Viñeta siguiente (Figura 3-4, área “C”): Espacio que muestra parcialmente el contenido de la siguiente viñeta a ser leída por el alumno. La intención es dar un aspecto visual más cercano a una tira cómica. Al final de cada sección esta área está vacía (Figura 3-1, área “C”).

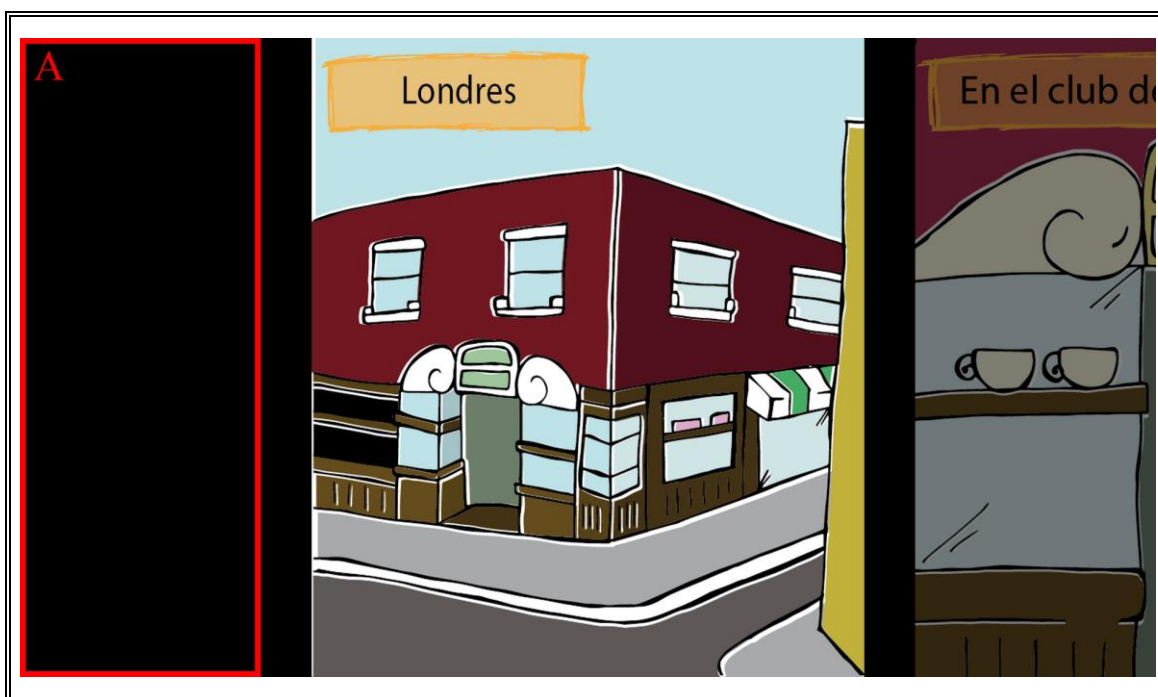


Figura 3-5: Vista de la pantalla de inicio de las secciones

Durante la actividad de lectura y al final de ésta, el profesor que utilice el instrumento puede acceder a un reporte del progreso de cada estudiante. Como se puede observar en

la Figura 3-6, este reporte muestra el porcentaje de acierto de cada estudiante en los respectivos niveles, el tiempo que le ha tomado la evaluación y si las respuestas que ha dado a cada pregunta son correctas o incorrectas.



Figura 3-6: Ejemplo de reporte de desempeño del instrumento

En este reporte, las respuestas dadas por el alumno a cada pregunta dan cuenta del nivel de dificultad en el que se encontraba al contestarlas, obteniéndose así un gráfico que muestra el desempeño del estudiante a lo largo de la sesión, y así su capacidad de comprensión lectora. En la Figura 3-6 observamos que el alumno Carmach contestó correctamente y se mantuvo en el nivel avanzado hasta la sección 9, donde se equivocó y por lo tanto el sistema le presentó nuevamente la sección, pero en un nivel de dificultad intermedio y con preguntas de este mismo nivel. En el ejemplo, el estudiante respondió correctamente y avanzó a la siguiente sección. Esta situación se repite en la sección 10, lo que deja al alumno en el nivel básico de lectura. En la sección 11 la situación es similar, pero dado que ya se encuentra en el nivel básico, sus respuestas se reflejan en una sola celda del reporte, primero un error y luego una respuesta correcta. Luego avanza en la historia, contestando correctamente y progresando al nivel avanzado hasta la sección 16, donde comete su último error antes de terminar la lectura finalizando en el nivel intermedio.

4. ARQUITECTURA DE SOFTWARE

Con la idea de lograr que el desarrollo fuera accesible desde distintos tipos de dispositivos, como computadores, tablets y teléfonos inteligentes, se diseñó el instrumento basándose en una arquitectura web cliente-servidor, particularmente una de tres capas. Así, como se puede ver en la Figura 4-1, el software fue dividido en la capas de presentación, negocio y datos.

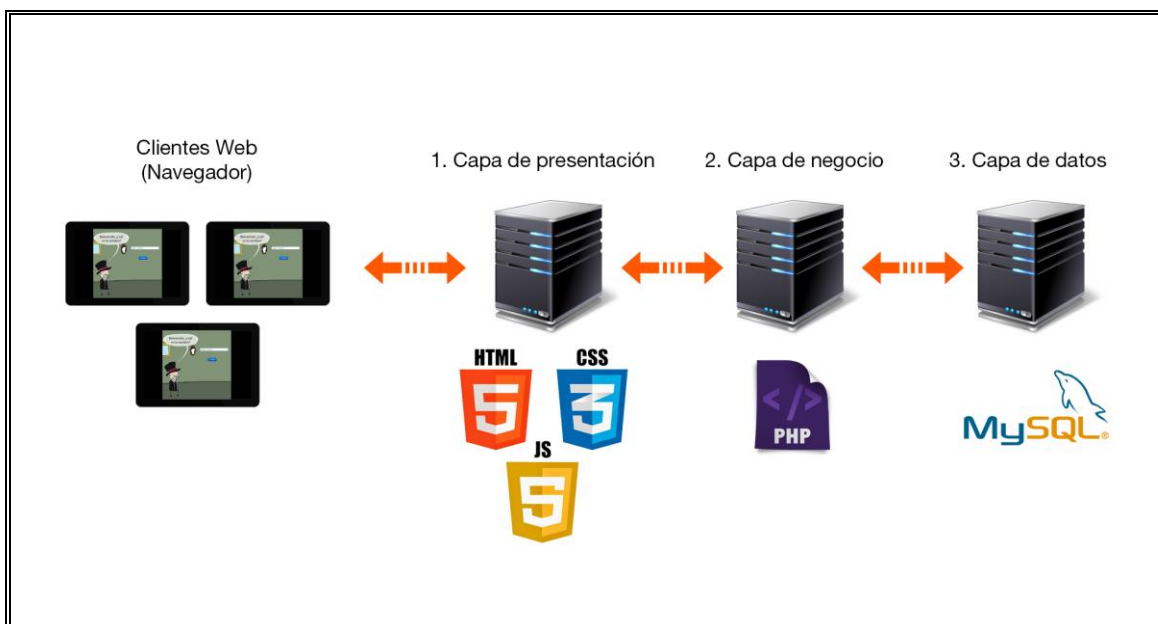


Figura 4-1: Arquitectura de software del instrumento

Cabe mencionar que si bien es factible que estas capas sean implementadas en distintos servidores y la conexión a ellos y entre ellos se haga por internet, en el caso del instrumento desarrollado las tres residían en el mismo equipo y el acceso a ellas se realizó a través de una red local.

Además de desligar el desarrollo del instrumento del dispositivo donde se iba a utilizar, esta arquitectura permitió separar la lógica de negocios de la lógica de diseño, es decir, separar el funcionamiento del instrumento de la forma como se ve el mismo. Así también, en caso de cualquier error se hace posible resolver el problema en la capa correspondiente, sin afectar el funcionamiento de las otras.

4.1 Capa de presentación

Esta es la capa del sistema con la que los alumnos interactúan, les presenta el instrumento y captura la información y las respuestas que estos van entregándole. Para su desarrollo se utilizaron tres tecnologías y lenguajes distintos, HTML, CSS y JavaScript, los cuales cumplen funciones determinadas dentro de la misma.

```

01. <!DOCTYPE html>
02. <html>
03.   <head>
04.     <title>Comics</title>
05.     <meta content="width=device-width, initial-scale=1.0, maximum-scale=1.0, user-
scalable=0;" name="viewport" />
06.     <meta name="apple-mobile-web-app-capable" content="yes" />
07.
08.     <link rel="stylesheet" href="css/bootstrap.css" >
09.     <link rel="stylesheet" href="css/bootstrap-responsive.css">
10.     <link rel="stylesheet" href="css/index.css">
11.
12.     <script type="text/javascript" src="js/jquery.js"></script>
13.     <script type="text/javascript" src="js/jquery.touchSwipe.js"></script>
14.     <script type="text/javascript" src="js/jquery-ui.js"></script>
15.
16.   </head>
17.
18.   <body>
19.
20.     <img class="foto_izq" src=""></img>
21.
22.     <div class="div_centro">
23.       <img class="foto_centro" src=""></img>
24.     </div>
25.
26.     <img class="foto_der" src=""></img>
27.
28.   </body>
29.
30. </html>

```

Figura 4-2: Estructura básica del código HTML del instrumento

El lenguaje HTML define la estructura básica de la capa de presentación y, por lo tanto, del sistema web desarrollado. En esta caso, como se ve en la Figura 4-2, el instrumento consta de tres secciones principales que son descritas dentro de los *tags body* del archivo, primero la que corresponde al área de la viñeta anterior (foto_izq), luego el área del contenido principal (foto_centro) y, finalmente, el área de la viñeta siguiente (foto_der). En este mismo documento se hace referencia a los archivos de estilo CSS, que configuran el aspecto de sistema, y a los archivos JavaScript que manejan la interacción con el usuario y van cambiando el contenido de las imágenes que se muestran.

El diseño de la interfaz del instrumento está descrito en lenguaje CSS y para su desarrollo se utilizó como base el *framework* Twitter Bootstrap, un conjunto de herramientas de código abierto para diseño de sitios y aplicaciones web desarrollado por Mark Otto y Jacob Thornton de Twitter. Este *framework* permitió al instrumento ser mucho más rico visualmente y atractivo para los alumnos, al incorporar botones de gran tamaño, colores llamativos y un aspecto moderno.

Finalmente la incorporación del lenguaje JavaScript, a través de la librería jQuery y sus derivados, permite al instrumento reaccionar a los gestos táctiles realizados sobre el dispositivo, de forma de que el alumno puede desplazarse a través de las distintas imágenes que componen cada sección de la historia como si se tratase de una serie de imágenes, completando así la apariencia y experiencia del cómic.

Cabe mencionar que la capa de presentación interactúa y se comunica únicamente con la capa de negocio del sistema desarrollado.

4.2 Capa de negocio

La capa de negocio es donde se establecen todas las reglas y la lógica que debe cumplir el instrumento. Se comunica con la capa de presentación y recibe las acciones y solicitudes de los usuarios, para luego procesarlas y presentar los resultados de acuerdo a estas reglas. Además, se conecta con la capa de datos, para solicitar al gestor de base de datos almacenar o recuperar datos de él.

En el instrumento desarrollado, esta capa está formada por múltiples archivos escritos en lenguaje PHP. Todos ellos basan su funcionamiento en la sesión del usuario actual, que se define cuando el alumno escoge su nombre desde una lista al comienzo de la historia, de esta forma se hace posible que los distintos archivos sigan al alumno en su avance por la misma y este se vaya guardando posteriormente.

Dentro de los archivos del instrumento destaca **comics.php**. Este es el archivo en el que se basa todo el funcionamiento de la historia que se muestra, se encarga de traspasar las imágenes de la sección actual a la capa de presentación y prepara la pregunta que el alumno deberá enfrentar al final de esta, definiendo los botones respectivos para poder contestarla.

Otro de los archivos destacables es **results.php**. Este, como su nombre lo indica, recibe la respuesta dada por el alumno y calcula si es correcta o incorrecta, de acuerdo a la información previamente almacenada en la base de datos. De acuerdo a esto y a los flujos definidos anteriormente en las Figuras 3-2 y 3-3, define una imagen y mensaje de *feedback* para el estudiante, el nivel de dificultad y la sección

a la que este pasa. Finalmente, este mismo archivo se encarga de registrar el progreso en la base de datos para que luego se puedan definir las estadísticas por alumno.

Cabe mencionar que cada vez que en la capa algún archivo necesitaba rescatar o guardar información en la base de datos, la conexión con la misma era supervisada y dirigida por el archivo **functions.php**. Este archivo contiene la información necesaria para establecer dicha conexión y de esta forma se favorece la modularidad del software y las responsabilidades de cada archivo.

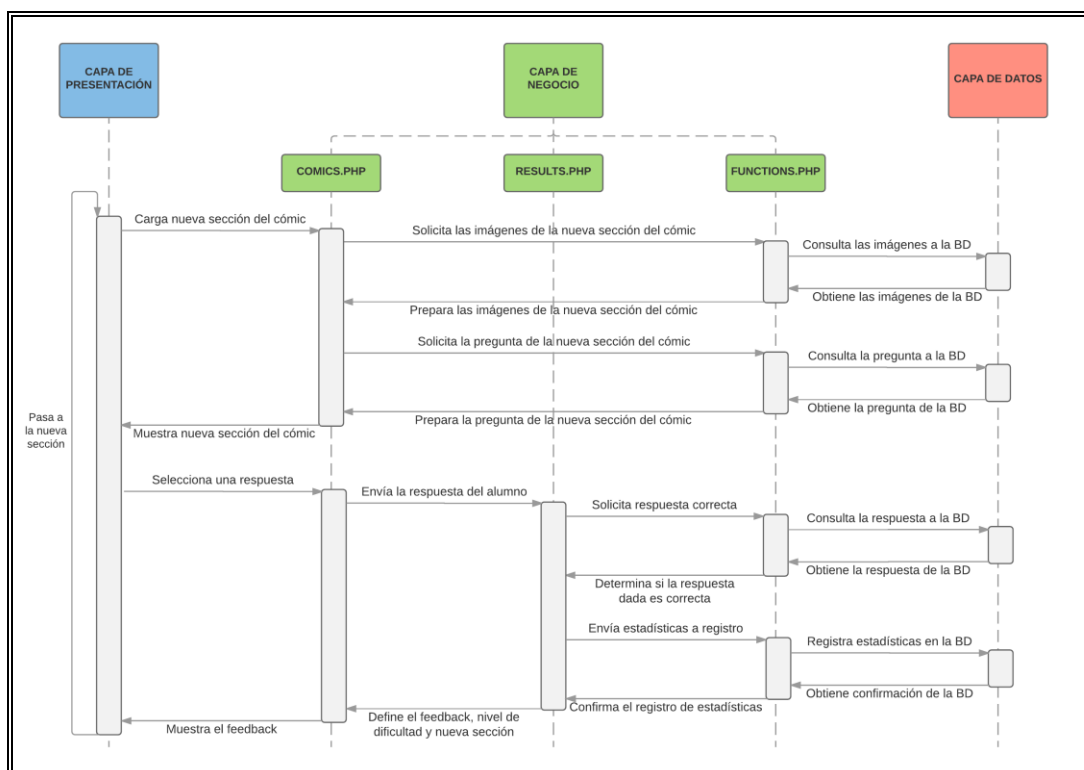


Figura 4-3: Diagrama de secuencia del funcionamiento de la capa de negocio

En la Figura 4-3 se presenta un diagrama de secuencia del funcionamiento de la capa de negocio, de forma que se entienda mejor el flujo de los datos y cómo el instrumento hacía el seguimiento del progreso de los alumnos a través de la historia.

4.3 Capa de datos

La capa de datos es donde reside toda la información y los datos asociados al instrumento, sus imágenes y preguntas, y a los usuarios, sus resultados y progreso. Está construida sobre el gestor de bases de datos relacionales MySQL, desarrollado bajo licencia dual GPL/Licencia comercial por Oracle Corporation y considerado como la base datos *open source* más popular del mundo, sobre todo para entornos de desarrollo web. Este se encarga de todo el almacenamiento de datos, recibir las solicitudes de almacenamiento o recuperación de información que vienen desde la capa de negocio. Esta capa está formada por cinco tablas distintas, que en conjunto guardan toda la información asociada al instrumento.

La primera de ellas, **frames**, contiene datos asociados a las imágenes de la historia, definiendo para cada sección y nivel de dificultad, el número total de imágenes a mostrar.

La tabla **questions**, como su nombre lo indica, mantiene información sobre las preguntas que se hacen al final de cada sección, definiendo para cada una de estas y nivel de dificultad, la respuesta correcta de la misma. Estas últimas son representadas por un número 1, 2 o 3, dependiendo si la respuesta correcta era A, B o C, respectivamente. Durante la etapa de confección de las preguntas del

instrumento se definieron preguntas alternativas para una misma sección y nivel de dificultad, por lo que se decidió incluir en esta tabla un valor **qNumber** para identificarlas.

La tabla **users**, mantiene una lista de nombres y apellidos de los alumnos que van a usar el software. Esta lista era facilitada previamente por los establecimientos, de forma que se podía llenar la tabla con la información indicada antes de la experimentación en colegios. El objetivo de guardar esta información era que los alumnos se identificaran rápidamente en el software al encoger sus nombres desde una lista que se desplegaba al comienzo del instrumento. Además, en ella se puede guardar la información de los emails y contraseñas para los usuarios administradores, datos con los cuales acceden a la plataforma de reportes. Para indicar el tipo de usuario y diferenciar a los administradores, por ejemplo los profesores de un establecimiento, se usa el campo **type** de la tabla.

La tabla **results** es la encargada de guardar las respuestas que los alumnos dan a lo largo del instrumento. Para esto hace un cruce de información entre las tablas **users** y **questions**, registrando el id del usuario y el id de la pregunta que se está respondiendo, así como la alternativa escogida por el alumno.

Finalmente, la tabla **progresses**, de manera similar a la anterior, va registrando el avance y las estadísticas acumuladas del alumno a lo largo de la historia. Al incorporar la información de las tablas **users** y **frames**, es capaz de registrar, para cada usuario, la sección y el nivel de dificultad en el que se encuentra. Esto permite, por ejemplo, recuperar la sesión de un alumno en caso que la conexión del dispositivo que estaba utilizando falle. Además, esta tabla mantiene la

información sobre la cantidad de tiempo acumulado que el alumno lleva interactuando con el instrumento y el número de respuestas correctas e incorrectas que ha dado.

En la Figura 4-4 se presenta un diagrama de las entidades de la base de datos, sus atributos y la forma en como estas se relacionan, de forma que se entienda mejor el modelo de datos detrás del instrumento.

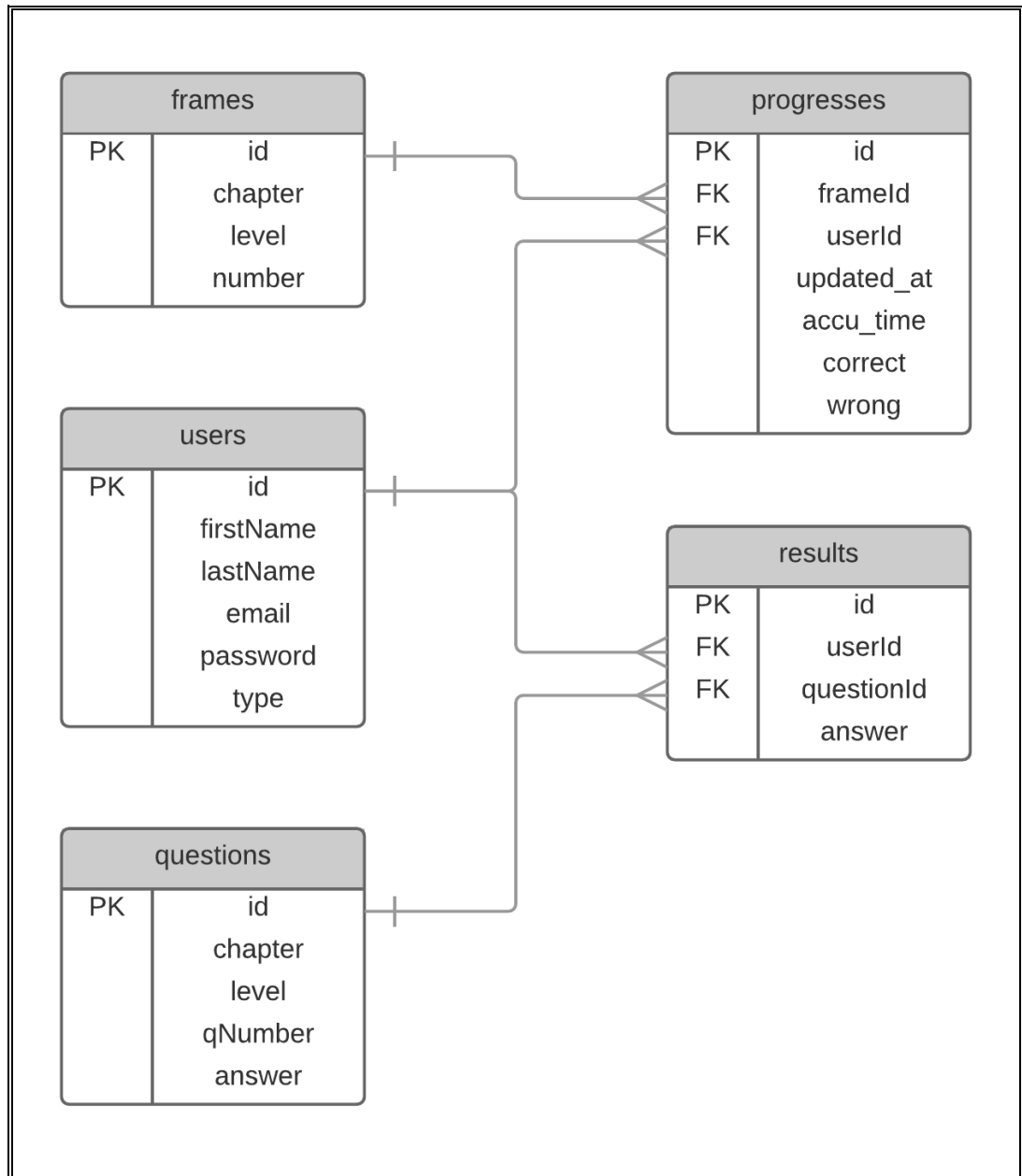


Figura 4-4: Diagrama de la base de datos del instrumento

5. EXPERIMENTACIÓN EN COLEGIO

Antes de iniciar la experimentación en los colegios, estos debieron ser seleccionados entre aquellos que aplican a sus estudiantes la prueba SEPA, de Lenguaje y Comunicación, alineada con el plan de estudios nacional y que proporciona información relacionada con el rendimiento y progreso del estudiante durante el año escolar (Alfonso et al., 2010). Esta selección se hizo con el objetivo de contar con colegios que evaluaran a sus estudiantes con una prueba estandarizada y, de esta forma, fuera posible medir la efectividad del instrumento desarrollado comparando los resultados obtenidos para establecer la validez de constructo del mismo. Dadas las diferencias entre ambos tipos de pruebas, estática (SEPA) y dinámica (instrumento desarrollado), las comparaciones se realizaron considerando los niveles de logro medidos en la prueba estática, tanto en las diversas áreas que evalúa como en el promedio de estas, con los resultados calculados por el instrumento. Adicionalmente a los resultados de SEPA, se utilizaron las calificaciones de los alumnos en la asignatura de Lenguaje y Comunicación para validar el instrumento.

La prueba SEPA se aplicó en forma independiente por los profesores de cada establecimiento, quienes debieron seguir un protocolo de aplicación proporcionado por el equipo de MIDE-UC. La duración de cada aplicación fue de 90 minutos y se tomó unos pocos meses (entre 1 y 2 meses) antes de la experiencia con el instrumento en estudio en cada colegio.

Para la experimentación se utilizaron tablets de 10.1". La ansiedad de uso de los dispositivos tecnológicos se observó sólo en aproximadamente un 20% los alumnos de cada curso. De manera generalizada, los alumnos manifestaron un gran interés y entusiasmo por la actividad realizada.

5.1 Primera iteración

Tal como se mencionó en el capítulo 2, durante la primera iteración se analizó el comportamiento de la herramienta con usuarios reales, con la intención de medir tanto la usabilidad del software como la dificultad general del instrumento desarrollado. Con este fin, se realizó una prueba en el Colegio Ecole Noel, establecimiento público de la Región Metropolitana, en la que se contó con alrededor de 30 alumnos de tercero básico.

Para llevar a cabo la prueba, se desarrolló el módulo de reporte de desempeño para el instrumento, mencionado anteriormente, con el cual se pudo observar el progreso de los alumnos durante la misma.

Una vez en el colegio, el sistema fue montado dentro de la biblioteca del establecimiento, con la intención de realizar la prueba en un lugar donde los estudiantes están acostumbrados a mantenerse en silencio (Figura 5-1). Luego de que los estudiantes fueron distribuidos aleatoriamente en los puestos, se realizó una pequeña introducción para explicar el funcionamiento del instrumento y se dio comienzo a la sesión de trabajo, con una duración aproximada de 45 minutos.

Durante la intervención los alumnos fueron acompañados por los investigadores y el profesor jefe del curso. Los primeros a cargo del buen funcionamiento del

instrumento y de guiar a los alumnos durante la sesión, y el segundo a cargo de mantener el orden dentro de la sala y cuidar que se siguieran las instrucciones.

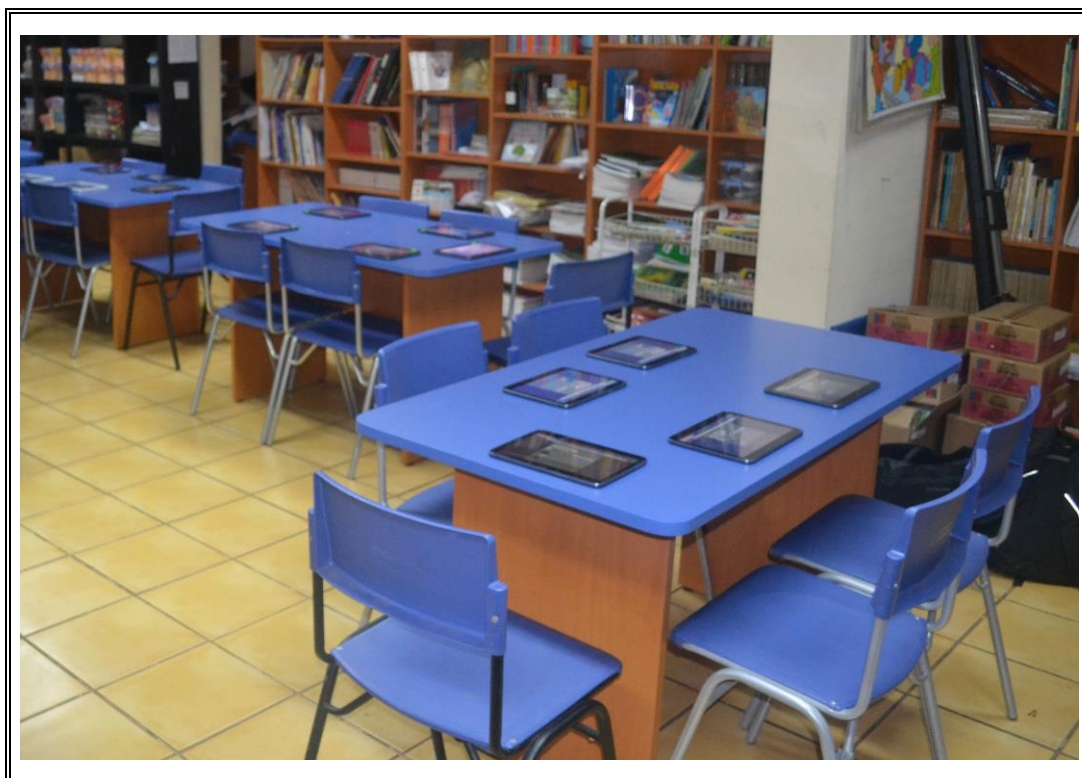


Figura 5-1: Distribución y montaje del instrumento durante la experimentación

5.2 Segunda iteración

Durante la segunda iteración, el instrumento se probó en dos colegios de la Región Metropolitana, uno del sector público, Escuela San José Obrero en la comuna de Peñalolén, y otro del sector privado, Colegio Monte Tabor y Nazaret en la comuna de Lo Barnechea, acumulando con esto una muestra de 182 alumnos (81 mujeres y 101 varones) de tercero básico de edades entre 9 y 10 años.

De forma similar a la primera iteración, los alumnos fueron distribuidos aleatoriamente en la sala de computación de cada establecimiento, donde se realizó la introducción y se procedió a la aplicación del instrumento (Figura 5-2), la cual, como puede verse en la Tabla 5-1, tardó aproximadamente 45 minutos en cada curso, demorando los alumnos entre 13 y 45 minutos. En cada uno de los cursos se evaluó el nivel de comprensión lectora de los alumnos utilizando el instrumento en estudio.



Figura 5-2: Aplicación del instrumento durante la experimentación

Tabla 5-1: Notas y tiempos obtenidos por los alumnos en el instrumento

Resultados	Notas Instrumento (1.0 a 7.0, 7.0 mejor valor)	Tiempo Instrumento (min)
Mínimo	2.20	13.22
Máximo	7.00	45.43
Promedio	4.65	25.02

El resultado de la prueba SEPA, las calificaciones de los alumnos en la asignatura de Lenguaje y Comunicación y el reporte de cada alumno obtenido a través del sistema en estudio, se utilizaron para realizar las validaciones del mismo (Anexo D). Las calificaciones que se analizaron corresponden al período en el cual se aplicaron los instrumentos.

6. RESULTADOS

A continuación se presentan los resultados obtenidos en ambas iteraciones de la experimentación en colegios. Además se entrega un análisis de estos y las principales consecuencias que tuvieron en el rediseño del instrumento.

6.1 Primera iteración

Una vez completada la prueba en el colegio se rescataron los resultados de los alumnos y se realizaron análisis descriptivos de los niveles de logro de estos para determinar, como se mencionó anteriormente, la usabilidad y dificultad general del instrumento.

Como resultado de estos análisis se obtuvo un Alfa de Cronbach de 0.829, considerado como un número adecuado, particularmente en casos de pruebas de contenidos disciplinares. Luego, como se ve en la Tabla 6-1, se determinaron distintos Alfa de Cronbach considerando que cada sección del instrumento era un ítem por separado y se calculó la correlación ítem-test para analizar si era necesario eliminar alguna de estas secciones. De este análisis destaca que la media total del instrumento no varía mucho si se elimina una sección en particular y el índice de confiabilidad tampoco varía. Por lo tanto, se concluye que no es recomendable eliminar ninguna de las secciones, manteniendo todas ellas para conservar el correcto desarrollo de la historia. De esta forma se pudo establecer la correcta consistencia interna del instrumento.

Tabla 6-1: Análisis del índice de confiabilidad por sección

Sección	Media de la escala si se elimina la sección	Varianza de la escala si se elimina la sección	Correlación sección - total corregida	Alfa de Cronbach si se elimina la sección
Sección 1	28.5025	84.756	0.206	0.830
Sección 2	28.1357	81.199	0.234	0.831
Sección 3	28.5377	76.815	0.408	0.822
Sección 4	28.3116	76.458	0.521	0.815
Sección 5	28.1206	77.400	0.486	0.817
Sección 6	27.8794	78.238	0.502	0.817
Sección 7	28.0302	77.151	0.504	0.816
Sección 8	28.1558	77.021	0.452	0.819
Sección 9	28.8090	78.731	0.375	0.823
Sección 10	28.7739	80.600	0.343	0.825
Sección 11	28.5075	77.170	0.425	0.821
Sección 12	29.0352	78.458	0.385	0.823
Sección 13	28.9749	76.025	0.522	0.815
Sección 14	28.6382	74.384	0.570	0.812
Sección 15	28.6181	74.419	0.559	0.812
Sección 16	28.6181	74.712	0.526	0.814
Sección 17	29.0955	80.218	0.251	0.832

6.2 Segunda iteración

Los datos obtenidos de la segunda iteración de la experimentación en colegios se utilizaron para establecer la validez externa del instrumento. Para esto se

realizaron regresiones lineales simples tomando en consideración dos tipos de datos a la vez. Con el fin de establecer correlaciones favorables se compararon, por separado, los resultados obtenidos por los alumnos a través del instrumento con los porcentajes de logro de la prueba SEPA y, luego, con las notas de Lenguaje y Comunicación.

Tabla 6-2: Regresiones lineales en función de los resultados del instrumento

Resultados del instrumento	β	P value	R²
Nota de Lenguaje y Comunicación	0.3	< 0.001	0.241
Porcentaje de logro en prueba SEPA	0.105	< 0.001	0.315

Como se puede observar en la Tabla 6-2, mediante dichas regresiones lineales simples se encontraron correlaciones favorables y positivas entre el resultado obtenido en el instrumento en estudio y las notas que tienen los mismos alumnos en la asignatura de Lenguaje y Comunicación, explicando el 24.1% (R cuadrado ajustado) de las variables que determinando dicho resultado. De la misma forma, se encontró una correlación positiva entre el resultado obtenido en el instrumento y los porcentajes de logro de los mismos alumnos en la prueba SEPA, explicando el 31.5% de las variables que determinando dicho resultado.

Con respecto a esto, los porcentajes obtenidos en ambos casos no son poco significativos, dado que el resultado académico es determinado por un largo número de variables de distintos niveles (nivel individual del alumno, nivel del colegio, nivel de la comunidad, etc.) (Taut et al., 2014).

Finalmente, con el fin de establecer un posible sesgo de género en el instrumento se llevó a cabo un análisis de los resultados de los estudiantes utilizando un test ANOVA de dos factores (two-way ANOVA) para determinar si existían diferencias significativas entre géneros. Como se observa en la Tabla 6-3, este test determinó diferencias significativas entre los promedios de hombres y mujeres en las medidas de todos los instrumentos. Esto es, en el instrumento de evaluación dinámica en estudio ($P = 0.0305$), en las notas de los alumnos en la asignatura de Lenguaje y Comunicación ($P = 0.00103$), y en los porcentajes de logro de la prueba estandarizada SEPA ($P = 0.0114$). Estos resultados son consistentes con los arrojados por la prueba nacional de 4° grado de Lenguaje (SIMCE) que también reporta diferencias a favor de las mujeres.

Tabla 6-3: Test ANOVA de dos factores para determinar sesgo de género

Instrumentos	F	gl	P value
Notas Lenguaje y Comunicación	11.13	1/180	0.00103
Nota Instrumento	4.757	1/180	0.0305
Tiempo (min)	3.56	1/180	0.0608
Porcentaje de logro SEPA	6.542	1/180	0.0114

7. CONCLUSIONES

Con base en los resultados presentados se puede concluir que el instrumento en estudio evaluó correctamente el nivel de comprensión lectora a través de las macro-habilidades para todos los integrantes de la muestra, independientemente del nivel inicial. Como el instrumento permite que cada alumno avance a su propio ritmo, se produjo una importante brecha entre quienes iniciaron en un nivel avanzado, y aquellos que iniciaron en un nivel básico, lo que se vio reflejado en el tiempo que le llevó a cada uno completar la prueba. A pesar de lo anterior, cabe destacar que en todos los casos se logró medir correctamente el nivel de comprensión lectora.

Durante las sesiones experimentales los alumnos mostraron signos de cansancio y falta de concentración después de 20 minutos. Considerando que en promedio les tomaba 25 minutos completar la prueba, se puede suponer por qué en las secciones finales los alumnos cometieron más errores. En vista de lo anterior, se recomienda mantener acotada la extensión de la historia para futuras aplicaciones.

Cada sesión se llevó a cabo por parte de los investigadores. Sin embargo el profesor de cada curso puede realizar la aplicación de la prueba, ya que el instrumento es accesible a través de un sitio web sin necesidad de ninguna configuración adicional.

Los resultados muestran, triangulando con distintos instrumentos, la capacidad predictiva del instrumento en estudio. Dados estos resultados, el instrumento dinámico y el estático estandarizado poseen una capacidad similar para evaluar el nivel de comprensión lectora. Sin embargo el instrumento dinámico promueve el desarrollo del

estudiante (Nazari & Mansouri, 2014) y da la posibilidad de medir y entender el nivel en el que se encuentra, y mejorar desde ahí su nivel. Adicionalmente, entrega resultados en tiempo real y tiene la factibilidad de hacerse accesible a un gran número de estudiantes a la vez a través de internet. Finalmente, al estar las preguntas insertas dentro de un cómic y gracias a la cercanía que éste produce con el lector (Cuadrado et al., 1999), debería disminuir el nerviosismo causado por las pruebas estandarizadas en los alumnos, lo que se traduciría en evaluaciones más confiables al haber menos variabilidad; queda como trabajo futuro validar esta hipótesis.

Los resultados de este estudio no son generalizables considerando cómo se eligió la muestra, el número de colegios evaluados y el rango de materias evaluadas. Se requiere una evaluación con una muestra más amplia elegida sin una intencionalidad y con una unidad curricular completa, para distintos años escolares. Finalmente, el uso de este tipo de instrumentos en aula facilita la Evaluación Formativa (Formative Assessment), utilizando las acciones del alumno y el error como oportunidad de aprendizaje, y orientando sobre las modificaciones de las actividades que puede desarrollar el profesor en aula (Black & William, 1998). Queda como trabajo futuro estudiar cómo se integra un instrumento como el presentado para el desarrollo de una unidad curricular bajo el concepto anterior.

BIBLIOGRAFIA

Alfonso, M., Santiago, A., & Bassi, M. (2010). An Alternative Pathway into Teaching: Placing Top University Graduates in Vulnerable Schools in Chile. Inter-American Development Bank.

Bannan-Ritland, B. (2003). The role of design in research: the integrative learning design framework. *Educational Researcher*, 32(1), 21-24.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in education*, 5(1), 7-74.

Calderón, J. F., Nussbaum, M., Carmach, I., Díaz, J. J., & Villalta, M. (2014). A single-display groupware collaborative language laboratory. *Interactive Learning Environments*, 1-26.

Centro de Medición de la Pontificia Universidad Católica de Chile (MIDE-UC). (2009a). Sistema de Evaluación de Progreso del Aprendizaje (SEPA). Obtenido de: <http://www.sepauc.cl/>.

Centro de Medición de la Pontificia Universidad Católica de Chile (MIDE-UC). (2009b). Informe de Instrumentos para Evaluación de Aprendizajes, Habilidades Intelectuales e Intrapersonales – Programa Enseña Chile. Unpublished manuscript, Inter-American Development Bank.

Centro de Medición de la Pontificia Universidad Católica de Chile (MIDE-UC). (2009c). Informe de Resultados de la Primera Aplicación de Pruebas SEPA – Enseña Chile. Unpublished manuscript, Inter-American Development Bank.

Centro de Microdatos, Departamento de Economía, Universidad de Chile. (2013). Segundo Estudio de Competencias Básicas de la Población Adulta 2013 y Comparación Chile 1998-2013. Obtenido de: http://www.educacion2020.cl/sites/default/files/estudio_competencias_final.pdf.

Centro Regional para el Fomento del Libro en América Latina y el Caribe. (2012). Comportamiento lector y hábitos de lectura. Obtenido de: http://cerlalc.org/wp-content/uploads/2013/03/fcbc1b_ComportamientoLector_Final.pdf.

Cerrolaza, O., Cuadrado, C., Díaz, A. & Martín, M. (1999). El placer de aprender. *Carabela* 41, 23-33. SGEL, Madrid.

Cimermanová, I. (2015). Using Comics with Novice EFL Readers to Develop Reading Literacy. *Procedia-Social and Behavioral Sciences*, 174, 2452-2459.

Cope, B., & Kalantzis, M. (2009). A grammar of multimodality. *The International Journal of Learning*, 16(4), 361-426.

Cook, T. D., Campbell, D. T., & Day, A. (1979). *Quasi-experimentation: Design & analysis issues for field settings* (Vol. 351). Boston: Houghton Mifflin.

Design-Base Research Collective (2003). Design-Based Research: An Emerging Paradigm for Educational Inquiry. *Educational Researcher* 32(1), 5-8.

Evans, J. (2013). From comics, graphic novels and picturebooks to fusion texts: a new kid on the block!. *Education 3-13*, 41(2), 233-248.

Farrell, J. (2015). Literacy Volunteer Organizations: Providing Alternative Education for Adults. *Delta Kappa Gamma Bulletin*, 81(4), 75.

Fletcher, J. M. (2006). Measuring reading comprehension. *Scientific Studies of Reading*, 10(3), 323-330.

Fundación La Fuente, Adimark GFK (2010). Chile y los libros 2010. Obtenido de: http://www.uchile.cl/documentos/strongi-chile-y-los-libros-2010-fundacion-la-fuentei_84318_11_2501.pdf.

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge University Press.

León, J. A., & Pérez, O. (2003). Taxonomías y tipos de inferencias. JA León (Ed.), *Conocimiento y Discurso claves para inferir y comprender*, 45-66.

Nazari, B., Mansouri, S. (2014). Dynamic assessment versus static assessment: A study of reading comprehension ability in Iranian EFL learners. *Journal of Language and Linguistic Studies*, 10(2), 134-156.

O'Reilly, T., Weeks, J., Sabatini, J., Halderman, L., Steinberg, J. (2014). Designing reading comprehension assessments for reading interventions: How a theoretically motivated assessment can serve as an outcome measure. *Educational Psychology Review*, 26(3), 403-424.

Organization for Economic Cooperation and Development (OECD). (2011). PISA in Focus 8. Obtenido de: <http://www.oecd.org/pisa/pisaproducts/pisainfocus/49184736.pdf>.

Organization for Economic Cooperation and Development (OECD). (2014). PISA 2012 Results: What Students Know and Can do: Student Performance in Mathematics, Reading and Science (Volume I). Obtenido de: <https://www.oecd.org/pisa/keyfindings/pisa-2012-results-volume-I.pdf>.

Roschelle, J., Rafanan, K., Estrella, G., Nussbaum, M., & Claro, S. (2010). From handheld collaborative tool to effective classroom module: embedding CSCL in a broader design framework. *Computers & Education*, 55(3), 1018-1026.

Taut, S., Valencia, E., Palacios, D., Santelices, M. V., Jiménez, D., & Manzi, J. (2014). Teacher performance and student learning: linking evidence from two national assessment programmes. *Assessment in Education: Principles, Policy & Practice*, (ahead-of-print), 1-22.

TERCE. (2015). Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación. Obtenido de:

<http://www.unesco.org/new/fileadmin/MULTIMEDIA/FIELD/Santiago/pdf/TERCE-Cuadernillo2-Logros-aprendizaje-WEB.pdf>.

Valerie, L., & Abed, F. (2013, January). Comic life + writing = motivated student writers: incorporating visual graphics to teach writing. In *Preparing Teachers to Teach Writing Using Technology* (pp. 163-178). ETC Press.

ANEXOS

ANEXO A: ARTÍCULO PUBLICADO POR LA REVISTA INTERACTIVE LEARNING ENVIRONMENTS



Interactive Learning Environments



ISSN: 1049-4820 (Print) 1744-5191 (Online) Journal homepage: <http://www.tandfonline.com/loi/nile20>

A single-display groupware collaborative language laboratory

Juan Felipe Calderón, Miguel Nussbaum, Ignacio Carmach, Juan Jaime Díaz & Marco Villalta

To cite this article: Juan Felipe Calderón, Miguel Nussbaum, Ignacio Carmach, Juan Jaime Díaz & Marco Villalta (2014): A single-display groupware collaborative language laboratory, Interactive Learning Environments, DOI: [10.1080/10494820.2014.917111](https://doi.org/10.1080/10494820.2014.917111)

To link to this article: <http://dx.doi.org/10.1080/10494820.2014.917111>



Published online: 21 May 2014.

A single-display groupware collaborative language laboratory

Juan Felipe Calderón^{a*}, Miguel Nussbaum^a, Ignacio Carmach^a, Juan Jaime Díaz^a and Marco Villalta^b

^aComputer Science Department, School of Engineering, Pontificia Universidad Católica de Chile, Santiago, Chile; ^bSchool of Humanities, Universidad de Santiago de Chile, Santiago, Chile

(Received 18 September 2013; final version received 16 February 2014)

Language learning tools have evolved to take into consideration new teaching models of collaboration and communication. While second language acquisition tasks have been taken online, the traditional language laboratory has remained unchanged. By continuing to follow its original configuration based on individual work, the language laboratory fails to take advantage of the potential provided by collaborative learning. We propose the use of a language laboratory based on single-display groupware (SDG) for learning vocabulary, grammar, pronunciation, and listening comprehension. By adopting SDG, the language laboratory benefits from the advantages of small group collaborative learning. In this paper, we first describe the design and pedagogical merit of the SDG collaborative language laboratory. We then share the results of a quasi-experimental pre–post comparison study, and use an observation guideline to analyze whether the conditions for collaborative learning have been fulfilled. Based on the results of this study, we conclude that developing pronunciation skills can be more effective when using a collaborative language laboratory versus an individual language laboratory. In addition to this, it can also be concluded that collaborative learning is most effective when accompanied by adequate instructional design.

Keywords: computer-assisted language learning; computer-supported collaborative learning; single-display groupware; language laboratory; speech recognition; speech synthesis

1. Introduction

Second language acquisition (SLA) has become increasingly relevant on a global level. This is due not only to the increased need for learning a new language, but also to the introduction of new technologies. These technologies allow for new forms of communication and interaction between students, both face-to-face and remotely via the internet (Rama, 2012). When SLA is supported by the use of computers in this way, it is known as Computer-Assisted Language Learning (CALL).

The history of CALL can be divided into three stages, with the technological developments of each stage currently in co-existence (Warschauer & Healey, 1998):

- Behaviorist CALL: focuses on learning a second language by repetitively completing exercises on an individual basis. Here, the computer acts as a tutor by checking answers and giving the corresponding feedback (Lee, 2000).
- Communicative CALL: focuses on fostering communicative situations where students must be capable of producing texts and generating dialogs. Examples include systems that allow for the reconstruction of texts, role playing, and video games (Bax, 2003).
- Integrative CALL: focuses on integrating four basic skills needed for language learning (reading, writing, listening, and speaking) in a single activity. Here, the computer goes beyond the role of tutor by coordinating the entire learning process. In turn, these four skills should each be associated with other areas or concepts of language learning, such as grammar, vocabulary, pronunciation, and cultural awareness (Levy, 2009).

Although the teaching models applied to SLA have evolved, this evolution is not always reflected in the development of SLA applications. One such application is the language laboratory (Roby, 2004).

Language laboratories initially focused on listening comprehension and pronunciation activities without the support of a computer (Morton, 1960; Harvey, 1978; Vanderplank, 2009). Although students could receive feedback on their work from a teacher, this was often without direct interaction between the two. When language laboratories were first introduced, students used individual cabins equipped with headphones and a microphone, as well as recording and playback devices. With the development of computer technology, this new technology was incorporated into language laboratories in several different ways. Visual support was added to help organize the activities, as well as supplementary audiovisual materials (Barr, Leakey, & Ranchoux, 2005; Pranita, 2010). Automatic speech recognition (ASR) software

was used to test pronunciation (Neri, Cucchiarini, & Strik, 2003; Xu & Seneff 2009). Synthesized voice was used to reinforce pronunciation (Handley, 2009). The internet was used to access materials and activities (Shingal, 1997; Hsu, 2005), and audiovisual recordings were used to create narratives (Wagener, 2006), among others. However, the incorporation of new technology into the language laboratory has not affected its pedagogical design. The reported evidence suggests that the role of the language laboratory should be focused on constant drilling and practice combined with interaction. This interaction can be achieved through individual and collaborative tasks monitored by the teacher (Vanderplank, 2009).

Socio-cultural theories of learning suggest that SLA technologies should promote the development of a methodology to foster communicative skills, in line with the following hypotheses (Nguyen, 2010):

- Student learning must focus on processes of interaction and collaboration that are relevant to the learner.
- SLA activities should encourage group work, role play, and projects that take the students beyond the confines of the classroom.
- The teacher is a facilitator and should monitor student learning in both cooperative and/or collaborative work.

Given the importance of communication and integrated language learning in the design of such systems, these hypotheses can be related to the “Communicative” and “Integrative” stages of CALL development. This is also evidenced by the fact that the computers, and not just the teachers, mediate communication between students. Applications designed with these hypotheses in mind have been shown to be effective when learning a second language and developing communication skills. However, they often fail to integrate the four main language learning skills, focusing instead mainly on oral communication (Yang, Gamble, & Tang, 2011). Furthermore, an analysis of the contribution made by collaborative laboratories versus individual laboratories is also missing. This therefore gives rise to our first research question: when learning a language through integrated practice of the four skills, what advantages does a collaborative laboratory hold over an individual laboratory?

Collaborative learning (Dillenbourg, 1999) allows for skills such as negotiation and individual responsibility to be developed, as well as building group knowledge. These have all been defined as *21st century skills* (ATC21S, 2012). Collaboration in SLA has been shown to allow peers to discuss comprehension, improve the quality of discourse and develop responsibility and independence in learners.

When collaborative learning is aided by technology, it is referred to as Computer Supported Collaborative Learning (CSCL). In this case, the technology allows the interaction between participants to be mediated. This is achieved by sharing information, administering homework assignments, establishing rules and roles, and facilitating the acquisition of new knowledge (Zurita & Nussbaum, 2004a). CSCL shares certain aspects

from the categories proposed in the taxonomy by Warschauer and Healey (1998), especially the incorporation of technology to encourage communication and interaction among peers. In language learning, this concept is defined as Computer Mediated Collaborative Learning (Warschauer, 1997; Yamada, 2009). This concept has also been adopted by Levy (2009) in developing language learning skills by reviewing Sykes' analysis (2005) of three types of synchronous group discussion: written chat, oral chat and face-to-face discussion.

In order for collaborative learning to be successful, certain conditions must be fulfilled regardless of whether or not they are mastered (Szewkis, Nussbaum, Rosen, Abalos, Denardin, Caballero, Tagle, & Alcoholado, 2011). These conditions include the existence of a common goal (Dillenbourg, 1999), positive interdependence between peers (Johnson & Johnson, 1999), coordination and communication between peers (Gutwin & Greenberg, 2004), individual accountability (Slavin, 1996), awareness of peer work (Janssen, Erkens, Kanselaar, & Jaspers, 2007), and joint rewards (Axelrod & Hamilton 1981).

In the field of language learning, some of these conditions have already been analyzed within the specific context of developing cooperative work among peers. In particular, this analysis refers to the incorporation of positive interdependence or personal accountability in activity design (AbuSeileek, 2012). In general terms, it is noted that this is achieved when there is less interference by peers. This shows that when developing communication and speaking skills, the contribution made by each individual should be accepted by their peers (AbuSeileek & AbuAlshar, 2012). Further conditions or aspects have also been shown to develop within collaborative work. These include the emergence of an *expert* among peers, discussions about how a task should be performed, how students work when faced with challenges, and how they develop interpersonal relationships (Leahy, 2008). The presence of these additional conditions or aspects allow learner behavior to be analyzed within the context of CALL. This therefore gives rise to our second research question: when learning a language in a collaborative, face-to-face activity, which are the most relevant conditions of collaboration to be fulfilled?

Incorporating collaboration into CALL tools requires an understanding of the processes in which the students are engaged during these activities (Hampel, 2009). An important aspect to review is the influence of technology on student behavior regarding collaborative activities (Leahy, 2012). One such case is the use of mobile platforms and sensors. In this case, technology helps to organize and mediate social interactions, regardless of the place and time in which the activities are performed (Ogata, 2008). In language learning, advantage can be taken of the ubiquity of these devices to encourage collaboration when completing activities with common goals, using personalised context-aware techniques in order to enhance learner learning interest and efficiency (Chen & Li, 2010). This should be accomplished by using the devices' multimedia capabilities to record videos and images as a group (Ogata & Yano, 2003). Another

example of that is virtual worlds, where real-life environments are generated in 3D (Shih & Yang, 2008) and simulated by the computer systems. In these virtual worlds, the participants are represented by avatars or representations of the users' identities (Li & Wong, 2010). It has been demonstrated that a virtual space for face-to-face interaction favors collaboration (Wang & Chen, 2010) and in particular the development of communication and peer support through voice and image interaction (Yamada, 2009). For example, Second Life provides a stimulating environment for learners to engage in a range of social interactions involving collaborative dialog (Peterson, 2012). The use of virtual worlds has also been shown to be effective in SLA at both utterance and discourse levels (Zheng, Young, Wagner, & Brewer, 2009), as well as the acquisition of communication skills (Berns, Palomo-Duarte, Doderer, & Valero-Franco, 2013), and the construction of meaning (Blake, 2011; Deutschmann, Panichi, & Molka-Danielsen, 2009). However, these models also have some disadvantages. One of the disadvantages is the high cost of implementation, given that they require one or more device per student. Furthermore, the computers are not shared, making it difficult for peers to be aware of each other's work. This is because each student's answers are only shown on their respective devices, both for the ubiquitous in-person activities as well as those in the virtual world. In the latter case, coordination and communication among peers can be compromised by connectivity and latency issues on the various different devices (Garrido-Iñigo & Rodríguez-Moreno, 2013). This situation is critical because the students require permanent connectivity in order to provide relevant and timely feedback and/or the possibility of mutual reinforcement between peers. Another concern is that the absence of face-to-face contact could result in less engagement by students. As students can only see the avatars and do not receive verbal or non-verbal cues from their fellow learners, this could result in problems with the interaction among peers (Macías-Díaz 2008b in Duncan, Miller, & Jiang, S 2012).

The SDG model has been used to develop collaborative learning (Stewart, Bederson, & Druin, 1998), which allows several students to learn collectively in front of a single screen (Infante, Hidalgo, Nussbaum, Alarcón, & Gottlieb, 2009). The information shown to users is shared on a single display device, with multiple input devices for shared control allowing the students to act simultaneously and in the same place (Kaplan, DoLenh, Bachour, Yi-ing, Gault, & Dillenbourg, 2009). One important aspect that promotes interaction between students is the fact that each student must work with their own objects on the screen using their own input device. This forces them to participate and play a central role in their own learning process (Infante et al. 2009). This interaction allows for the emergence of a shared interaction pattern, leading to the development of better quality discussions in environments where face-to-face interactions do not take place (Chung, Lee, & Liu, 2012).

In this study we propose the implementation of a language laboratory that uses small group collaborative learning as a teaching method. The laboratory is based on SDG, with the aim of studying the contribution of collaboration in the language laboratory and determining how the conditions of collaboration are fulfilled. First, we

describe the design of the pedagogical activity, detailing the skills to be developed by the students. Next, we describe the experimental design and detail the result of the experiment, along with the corresponding statistical analysis. This is followed by a discussion of the analysis, where the conditions for collaborative learning are reviewed. Finally, we present the conclusions.

2. Collaborative Language Laboratory

Kessler and Bikowski (2010) highlight that it is possible to observe the following forms of collaboration in SLA activities:

- Joint Collaboration: individuals should have equal responsibilities.
- Parallel Collaboration: individuals should have different responsibilities, but work towards the same objective.
- Incidental Collaboration: individuals collaborate based on the requirements that come up as they are carrying out the task.

This section presents the collaborative language laboratory design, which uses collaboration as its guiding principle. This was chosen as the guiding principle as it aims to give all students the same level of responsibility while working towards a common goal.

The design of the laboratory should take into consideration not only the students who will use the laboratory, but also the teachers. The teacher's role in this case is defined by orchestrations (Nussbaum, Dillenbourg, Dimitriadis, & Roschelle, 2013). These consist of previously-prepared lesson plans that integrate conventional and digital resources as well as combining the teacher's work with the students' laboratory work. In the practical laboratory work, the teacher's role includes explaining how to use the technological platform and offering instruction based on any difficulties faced by the students.

Table 1. Skills taught and material created for the collaborative language laboratory.

Skill	Number of Sessions	Number of Activities per Skill
Grammar	4	40
Vocabulary	4	42
Listening	4	27
Pronunciation	Integrated into the previous skills	Integrated into the previous skills

This project teaches the following skills: grammar, vocabulary, listening comprehension, and pronunciation (Table 1). These skills were taken from the curriculum defined by the Ministry of Education (Mineduc, 2010). The aim of this subject is for students to learn English and be able to use it as a tool to communicate on a basic level in a range of situations. In order to achieve this, these skills are combined in such a way so as to allow the students to acquire the knowledge that they need in order to obtain information and develop communicative functions. So as to incorporate these skills, and by following the aims of the ministry, orchestrations were developed and defined by a script that combines collaborative learning with the rest of the students' and teacher's activities.

The groups comprise three students as this has been shown to be the optimum number for collaborative learning in SLA (Hsu, Hsu, He & Chang, 2009). Each student has their own headset, which they can use to communicate with the system and listen to the system's instructions (Figure 1). The collaborative learning is based on ordering a series of elements (Zurita & Nussbaum, 2004), where the students must work together to build an ordered sequence using the elements belonging to each individual. Each member of the group can see the sequence and everyone else's personal elements using the shared screen. The sequence to be put in order will depend on the skills being developed:

- Grammar: a sentence must be constructed based on the words belonging to each student (Figure 2).
- Vocabulary: three photographs are presented, one per student, which must be matched with a word belonging to each of the students (Figure 3).
- Listening: a sequence of words must be built based on the order in which they appear in a text, which is listened to by the three members of the group simultaneously (Figure 4).

So that students will learn to accurately pronounce the words, they must select the words for each of the exercises by enunciating them. In order to build each sequence of words, each student must correctly pronounce the words that belong to them so that they are included in the sequence. For example, for the grammar exercise in Figure 2, an example is shown where the student correctly pronounces the first word of the displayed sentence. When this happens, the word disappears from the student's workspace and appears in the shared space. If the students build a sequence incorrectly, the system will indicate that there is an error and make them repeat the whole exercise by reassigning the words.

The model used to develop student feedback must be simple, clear, and in line with the design of the activity in order for it to be effective (Hémard, 1997). Thus, when a student makes a mistake in their pronunciation, the feedback given by the system is based on the model proposed by Mackey (2006) for interaction feedback. In this model, an expert repeats the concept or phrase attempted by the student so that the student may

repeat it correctly. Alternatively, the student can indicate that they do not understand, and the phrase will be repeated again.

In our case, the system acts as an expert by reproducing a recording of the word spoken by the student to the whole group. It then proceeds to give feedback to every member of the group through their headsets. When the student's pronunciation is similar enough for the word to be recognized but is still incorrect, the system offers feedback by repeating the correct pronunciation of the word (Figure 5). If the pronunciation is too dissimilar for the word to be recognized, the system will indicate that the word is incorrect. When the other students in the group receive this feedback, they can help their classmate to correct their pronunciation in a next opportunity. This feedback design therefore resembles incidental collaboration as it relates to an emerging situation which can be addressed by the students using collaboration.

The Automatic Speech-Recognition (ASR) system is used to evaluate pronunciation, with a voice synthesizer used for feedback. Both of these are provided by the Speech API (SAPI) version 5.4 Recognition and Synthesis libraries, which work on a Microsoft Windows operating system. Chen (2011) suggests that this library, with the corresponding modifications, is a free yet powerful tool that can be used to train oral skills in second language students.

Once the exercise is completed by the students, the system verifies that the sequence is correct. In this case, positive feedback is given and the students continue on to the next exercise (Figure 3, right). If the exercise was completed incorrectly, negative feedback is given and the same exercise is repeated (Figure 4, right). Students can only continue on to the next exercise if the sequence is correct, regardless of how many attempts they make.



Figure 1. In the foreground, students are using the collaborative language laboratory. In the background, students are using the individual language laboratory.

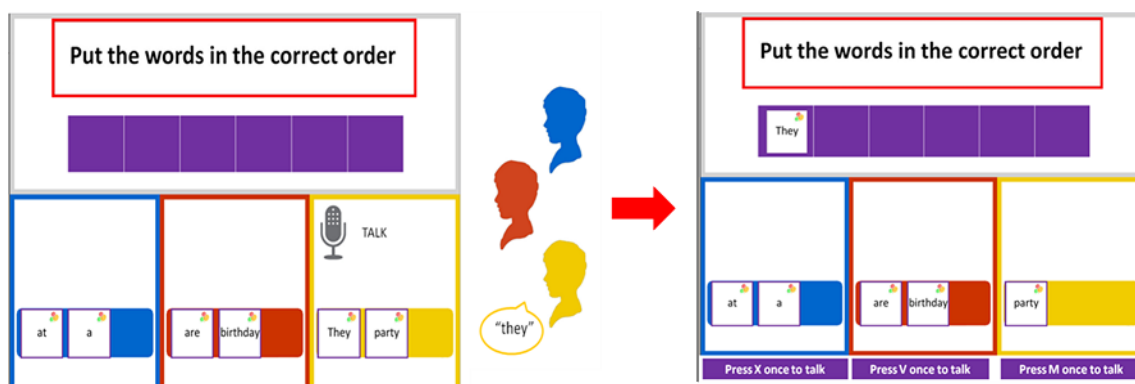


Figure 2. Example of exercises from a grammar activity. When a student pronounces a word correctly, it is moved to the shared space where the sequence is constructed. At the end of the exercise, the system confirms that the sequence is in the correct order.

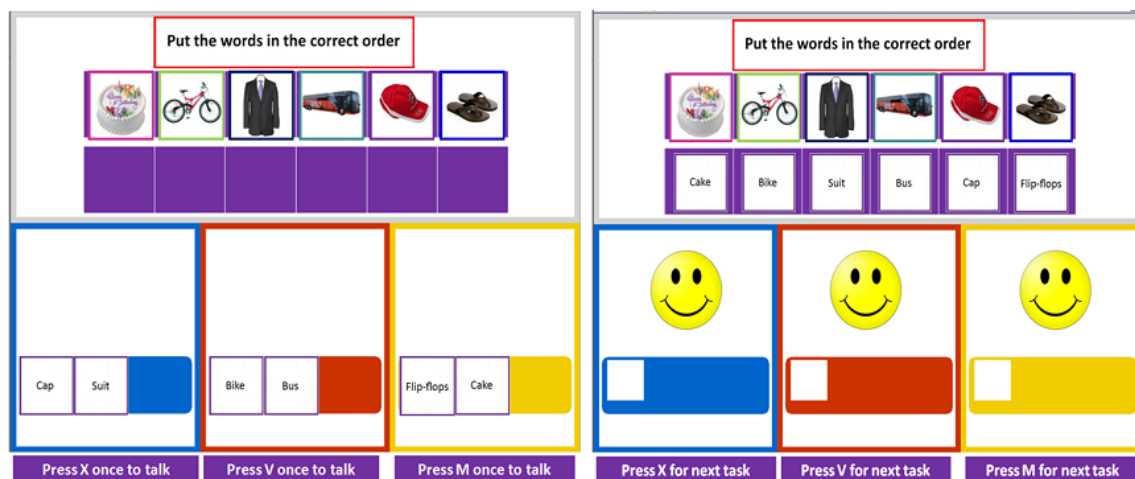


Figure 3. (Left) Screenshot of an exercise from a vocabulary activity. Students must match the pictures with words by pronouncing them. (Right) The system provides feedback; in this case the sequence is correct.

Figure 4. (Left) Screenshot of an exercise from a listening activity. Students must put the words in order according to their appearance in the text they are listening to. (Right) The system provides feedback; in this case the sequence is incorrect.

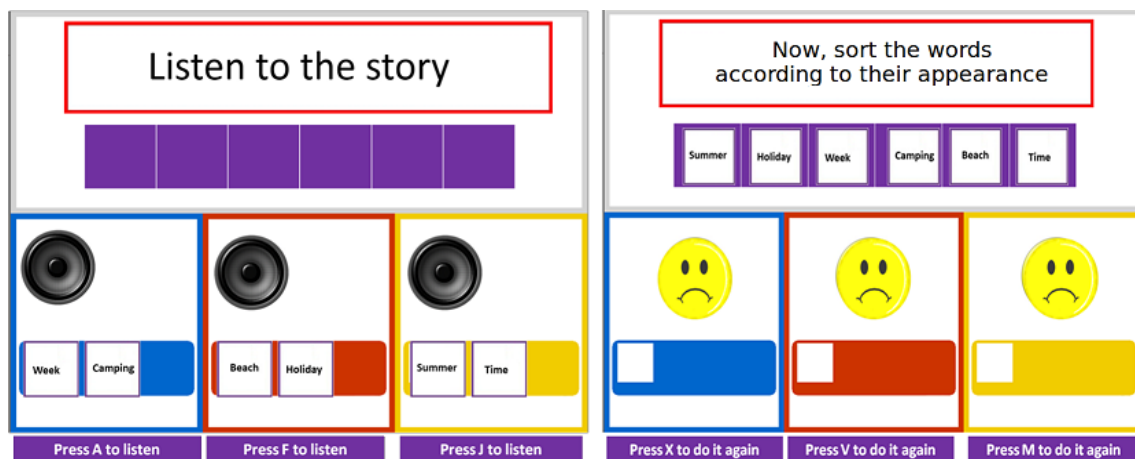


Figure 5. Sample screen of a word pronounced incorrectly, and the corresponding feedback given during a grammar activity.

3. Experimental Design

3.1 Tools used

One of the questions that this study looks to answer is: when learning a language through integrated practice of the four skills, what advantages does a collaborative laboratory hold over an individual laboratory? Therefore, to carry out this study, a second version of the collaborative language laboratory described in section 2 was built, with the same activities and technological resources, but adapted for individual work. The main difference is that there is only one set of elements, all of which belong to the student seated in front of the screen (Figure 6). With this, the student must build the sequence on their own without collaboration.

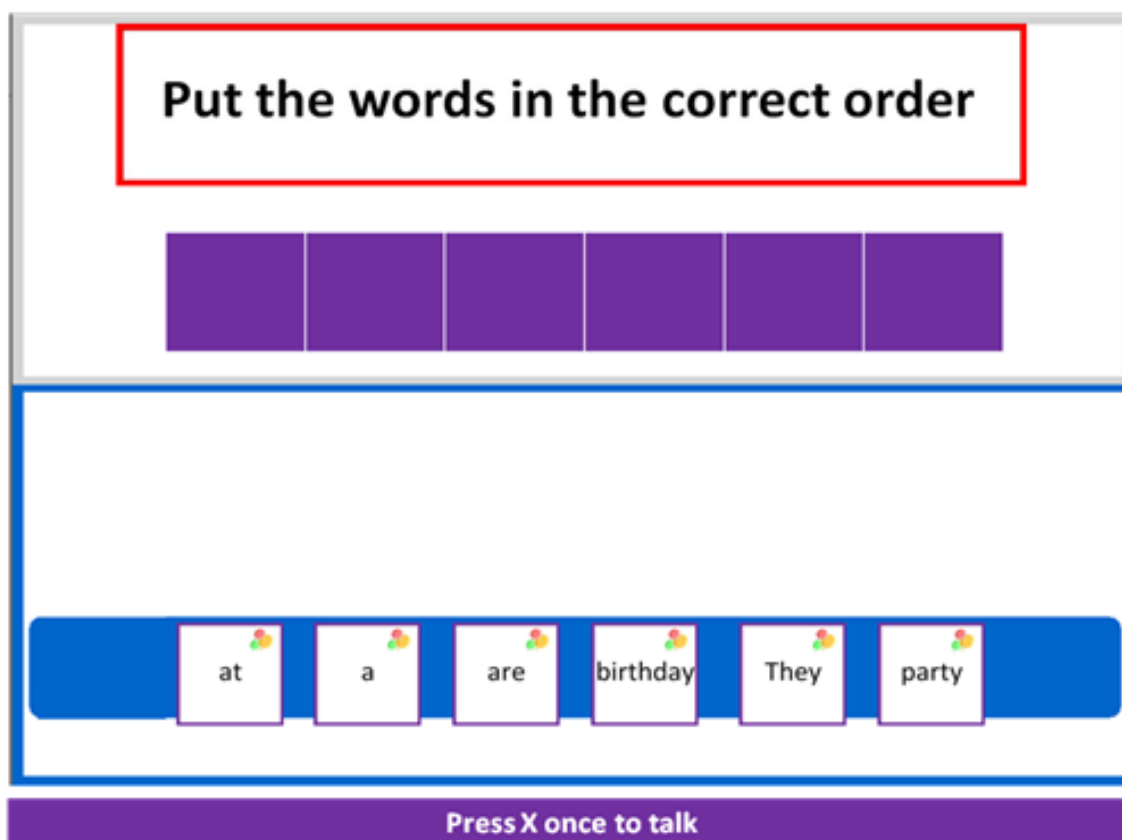


Figure 6. Sample screen of a grammar exercise in an individual language laboratory.

3.2 Study participants and procedure

For this study, an investigation was carried out with a quasi-experimental, pre-post type design. This was because the various groups that were to be analyzed were not necessarily even. The sample was composed of 6th grade students from a state-subsidized elementary school, whose ages range from 11 to 13. Three groups were defined and selected at random:

- Group working without the use of technology (control) (N=20: 8 boys and 12 girls)
- Group working in a collaborative language laboratory (N=24: 11 boys and 13 girls)
- Group working in an individual language laboratory (N=15: 8 boys and 7 girls)

Samples with small groups have already been used in other studies of this kind, where learning results are measured using quantitative variables and learner behavior observed using qualitative variables (Lee, 2011). The three groups were taught by the same teacher, who worked by following the orchestrations provided for this study (see Appendix 1). On the one hand, these orchestrations were designed to ensure that the teacher followed the curricular objectives set out by the Ministry of Education. On the other hand, they guaranteed that the teacher had the necessary resources to carry out the activity, ensuring that their use followed the scope and sequence proposed by the Ministry. The teacher plays a central role during the three phases of the orchestration. Each phase and the respective activities are assigned an initial quantity of time *a priori*. In the first phase, the professor introduces the activities to be carried out by the students using an activity that is familiar to everyone. The aim of doing so is to activate prior knowledge and include any additional necessary explanations. The activity in the second phase incorporates the use of language laboratories, where the teacher moves between groups to supervise the students' work and answer any questions that come up during the task. Finally, in the third phase, the teacher does an end-of-class activity where student learning outcomes are reviewed. The teacher also answers questions while the students work and is able to make certain decisions according to what comes up during class. The teacher may also change certain aspects of the orchestrations in order to adapt them to the specific needs observed during the class.

The role of the technological platform is to evaluate the development of language skills and it is designed to encourage collaboration. This is complementary to the role of the teacher since it relies on the development of an activity that involves many students, allowing them to work in a coordinated manner through a system of hardware and software. Furthermore, the use of multimedia resources serves as a learning aid for all students, specifically in this case through the visual display of material and the use of a speech recognition/speech synthesis engine. As mentioned previously, the use of technology does not necessarily lead to an improvement in learning if it is not

accompanied by an adequate design where both the teacher and students are well coordinated.

The students that worked without technology followed the same contents as the other two groups and worked on tasks that aimed to meet the same objectives. In this case, the teacher played a leading role by coordinating the interaction among students by following the same orchestrations as those for the groups using technology (Appendix 1). However, the orchestrations were adapted so that they could be followed without the use of language laboratories. In order to do this, sequencing activities were incorporated that students solved individually on paper, while pronunciation practice was carried out with the whole class.

3.3 Pre- and post-tests

The pre-post instrument consisted of a test developed by an expert, which students had to answer by using a computer platform. The test was multiple choice and included questions that independently evaluated the following four skills: vocabulary, grammar, listening, and pronunciation.

To evaluate vocabulary, grammar, and listening, the items on the test followed a similar design to that used in the activities included in the orchestrations (see Appendix 2). In order to measure the level of pronunciation, the same ASR engine from the language laboratory was used on the test. The items on the test evaluated the same skills and contents that the students developed in the activities. Table 2 details the number of items associated with each unit and skill.

The pre- and post-test items were selected using an item response analysis (Wright & Stone 1979). The items selected in this analysis allowed for them to be effectively discriminated (value between 0.3 and 0.99, in a range between 0 and 0.99, where a higher value means a higher degree of item discrimination).

Table 2. Skills, units covered, and number of items associated in the pre-post test.

Units	Vocabulary	Grammar	Listening	Pronunciation
Family and friends	5	6	2	7
Socializing	5	5	4	6
Celebrations	5	4	5	7
Holidays	5	5	3	7
Total	20	20	14	27

There were more items for pronunciation and fewer for listening because of the time it took to evaluate each one. In the case of pronunciation, each item corresponded to the evaluation of one word. The listening items also included a recording of a spoken text that students had to listen to (see Appendix 2).

Analysis of the reliability of the pre-post instrument was conducted using Cronbach's Alpha, calculated for each of the tests that measure the various language skills (Table 3).

Table 3. Cronbach's Alpha for the instrument used.

Skill	Cronbach's Alpha
Vocabulary	0.77
Grammar	0.71
Listening	0.89
Pronunciation	0.76

Fourteen sessions were held over a period of three months, in a computer room where both collaborative and individual laboratories were set up simultaneously (Figure 1).

3.4 Collaborative learning evaluation

The second objective of this study is to analyze how collaborative learning conditions are fulfilled by the students. In order to do so, classroom and video observations were carried out during 5 of the 14 sessions. These corresponded to the 1st, 4th, 7th, 10th, and 14th sessions, with the aim of observing how the variables that were analyzed evolved over time. In the field of CALL technology, various aspects have been analyzed regarding the interaction between learners, student attitudes towards the activity, motivation (Saggara & Zapata, 2008; Merisuo-Storm, 2007), development of communication in a face-to-face environment (AbuSeileek, 2012), and the quality of brainstorming when solving exercises (Leahy, 2008). To evaluate these aspects, an observation guideline was defined based on one that had been previously used in a study of collaborative SDG tools (Infante et. al 2009). The purpose of this guideline was to analyze the extent to which conditions for collaboration are fulfilled by incorporating the

specific aspects of language learning described here. The aspects to be analyzed were the following:

- **Communication:** measuring the level of communication in a face-to-face setting (AbuSeileek, 2012). This is measured using the number of person-to-person dialogs, person-to-group dialogs, and the number of times that students ask for and receive help from the group regarding the solution to an exercise. This also quantifies the number of times that solutions were imposed by a group member and not taken on board, but acknowledged by the rest of the group without further discussion.
- **Interaction:** observing aspects of the students' interaction, attitudes, and motivation (Saggara & Zapata, 2008; Merisuo-Storm, 2007). This is measured on a scale of 1 to 3, and sub-categorized as follows:
 - Positive interdependence: students feel that they are responsible for their own learning and that of their classmates.
 - Mutual trust: students trust each other; they do not question other group members' opinions, and they feel comfortable expressing their own.
 - Acceptance and tolerance: students are capable of accepting the opinions of other group members with whom they do not agree.
 - Motivation and interest: there is interest and motivation to work as a group to solve the problems in the activity.
- **Coordination:** observing aspects related to the students' attitudes towards the activity (Saggara & Zapata, 2008; Merisuo-Storm, 2007), and the quality of brainstorming when solving exercises (Leahy, 2008). This is measured on a scale of 1 to 3, and sub-categorized as follows:
 - Disciplined work: the established set of rules and roles are followed, with students working together as a group.
 - Requested support: support for performing individual or group activities is requested from people outside the group. A high score indicates that little support was requested from people outside the group; a low score indicates that a lot of support was requested.
 - Quality of brainstorming: students organize themselves to answer each exercise, with responses based not just on intuition but rather as the result of group planning.
- **Appropriation:** observing the students' attitudes toward the activity (Saggara & Zapata, 2008; Merisuo-Storm, 2007). This is measured on a scale of 1 to 3, and sub-categorized as follows:
 - Suitable handling of material: students master the use of the system (hardware and software).
 - Behavior towards the system: students develop adequate behavior towards the activity and its technological elements, without any discipline issues.

In order to apply this instrument, five observers were present (one per group) who had received previous training in how to apply the observation guidelines to a laboratory situation. Furthermore, videos were recorded for subsequent group analysis by the observers. This allowed the observations to be validated and for aspects that had not been considered by all the observers to be completed. The videos made it possible to agree on a set of criteria when applying the respective observation guidelines (see Appendix 3).

4. Results

4.1 Pre-post test results

To analyze the results from the application of the pre-post test, a two-tailed, unequal variance T-test was used to measure the significant differences between the different groups. An analysis of covariance (ANCOVA) was also conducted with the goal of adjusting the sample results based on a pre-test. The post-test was used as a way of discerning the differences among the samples. Cohen's d was used to measure the impact or effect size of a given group.

Table 4 shows the application of the T-Test on the samples, and Table 5 shows the significant differences post ANCOVA between groups ($p\text{-value} < 0.05$), along with the effect size using Cohen's d.

Significant differences ($p\text{-value} < 0.05$) can be observed for each skill and in every group between the pre- and post-tests (Table 4). The only exception is for pronunciation in the group using the individual language laboratory.

The significance of the results when comparing groups are indicated by the p-values and Cohen's d (Table 5). The three significant differences between the groups ($p < 0.05$) are highlighted in grey. The collaborative learning group stands out as improving pronunciation when compared to the other two, and for listening when compared to the control group. The negative values of Cohen's d for the technological groups in terms of vocabulary indicate that the technological groups performed worse than the control group. However, this difference is not significant.

Table 4. T-Test of the samples

		Pre test				Post test			
		Vocab ulary	Gramma r	Listenin g	Pronunci ation	Vocab ulary	Gramma r	Listenin g	Pronunci ation
Control N = 24	Min. value	3	0	3	1	5	0	3	1
	Max. value	19	9	11	16	19	9	12	15
	Average	8.45	2.45	6.65	9.2	11.15	3.45	7.35	10.05
	Std. Dev.	3.71	2.52	2.25	3.71	2.70	2.54	2.64	4.02
	p-value	--	--	--	--	<0.01	0.02	0.03	0.04
Collaborative N = 20	Min. value	2	0	1	2	4	1	2	6
	Max. value	13	15	14	15	17	18	14	17
	Average	7.21	2.58	6.67	9.71	9.58	4.83	9.33	12.67
	Std. Dev.	2.98	3.57	3.25	3.80	3.55	3.93	2.88	2.97
	p-value	--	--	--	--	<0.01	<0.01	<0.01	<0.01
Individual N = 15	Min. value	3	0	4	4	6	0	4	5
	Max. value	14	12	11	15	17	12	14	17
	Average	8.47	3.40	6.73	9.13	10.80	4.47	9.20	10.13
	Std. Dev.	2.88	3.16	2.52	3.20	3.00	3.16	3.36	3.89
	p-value	--	--	--	--	<0.01	0.01	<0.01	0.14

Table 5. Significant differences between samples (post-ANCOVA), and effect size using Cohen's d.

	Vocabulary		Grammar		Listening		Pronunciation	
	p-value	d	p-value	d	p-value	d	p-value	d
Collaborative v/s Control	0.11	-0.23	0.18	0.38	0.02	0.70	0.02	0.92
Collaborative v/s Individual	0.28	-0.11	0.76	0.31	0.90	0.06	0.03	0.62
Individual v/s Control	0.72	-0.13	0.30	0.05	0.08	0.76	0.95	0.07

4.2 Collaborative learning results

Table 6 shows the results from the analysis of the conditions required for collaborative learning in the collaborative group.

From the quantitative data in Table 6, it is evident that person-to-person and person-to-group dialogs are consistently present throughout the sessions, with a standard deviation of 5.61 and 2.98 respectively. This indicates that there is a greater dispersion in person-to-person dialogs. We can also see that the number of person-to-person dialogs always exceeds the number of person-to-group dialogs, something which is statistically significant (p-value of < 0.0003). This shows that the platform favors interactions that include the whole group over person-to-person interactions. It can also be noted that in all of the sessions, students were observed receiving more support than they requested; although this is not statistically significant (p-value > 0.05). Based on this, we can conclude that while students do not always necessarily receive more support than they ask for, it can be seen as a general trend. This could be attributed to the fact that the students are willing to support the rest of their peers proactively, without waiting for a specific request for help. On the other hand, the number of peer-imposed solutions from one peer to another decreased across the sessions, with the exception of the 4th session. Although this decrease was not significant (p-value > 0.05), the trend would suggest an improvement in communication among students.

Table 6. Results from the collaborative learning observations.

	Session	1 st		4 th		7 th		10 th		14 th	
	Type of activity	Vocabulary		Grammar		Listening		Vocabulary		Listening	
	Duration	24 min.		29 min.		30 min.		33 min.		28 min.	
	Number of questions	4		4		5		6		5	
	Total number of words	12		25		15		18		15	
Category	Variable	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.
Communication	Person to person	31.25	7.36	43.50	9.37	34.00	6.74	32.63	6.99	28.88	6.49
	Person to group	14.75	2.43	21.88	5.33	14.88	2.90	15.50	2.45	16.25	2.60
	Receive support	5.63	1.30	13.00	6.97	5.75	2.25	5.50	2.14	5.38	2.33
	Request for support	3.88	1.73	8.50	6.07	3.75	2.43	3.88	2.42	3.75	2.49
	Peer-imposed solutions	4.13	2.42	9.25	3.99	3.63	2.56	3.50	2.20	3.13	1.81
Interaction	Positive interdependence	2.25	0.46	2.38	0.52	2.50	0.53	2.63	0.52	2.63	0.52
	Mutual trust	2.63	0.74	2.13	0.35	2.63	0.52	2.75	0.46	2.75	0.46

	Acceptance and tolerance	2.25	0.46	2.13	0.35	2.50	0.53	2.75	0.46	2.75	0.46
	Motivation and interest	2.00	0.76	1.75	0.89	2.25	0.46	2.38	0.52	2.38	0.52
Coordination	Disciplined work	1.88	0.64	2.00	1.07	2.13	0.83	2.25	0.46	2.50	0.53
	Requested support	2.13	0.35	2.00	0.53	2.25	0.71	2.25	0.71	2.25	0.71
	Quality of brainstorming	1.75	0.46	2.13	0.83	2.25	0.71	2.25	0.71	2.25	0.46
	Suitable handling of material	1.25	0.46	2.38	0.52	2.38	0.52	2.50	0.53	2.75	0.46
Appropriation	Behavior towards the system	2.13	0.83	2.75	0.46	2.75	0.46	2.75	0.46	2.75	0.46

With regards to the qualitative data, the results for the majority of the categories that were analyzed are consistently closer to the maximum value than to the minimum (avg. = 2.36, s.d. = 0.21). In general terms, the variables indicate that these attributes are developed among the students. There is also a session where we can begin to observe a trend in each of these behaviors, whether this trend be an increase, decrease or constant. As this changes from variable to variable, the students' behavior over the observed sessions is not necessarily consistent, as discussed below.

Of the variables that were observed, mutual trust had the highest, significant value that was constant across the sessions, (p-value = 0.027), with the exception of the 4th session. This indicates that mutual trust is a characteristic which is constantly favored and that generally does not vary over time. The 4th session corresponded to a grammar activity where there was a notably higher number of peer-imposed solutions (Table 6). This can be aligned with a lower level of mutual trust given that the students tried harder to impose their opinions than in the other sessions and failed to acknowledge input from other group members.

Acceptance and tolerance and positive interdependence do not become constant until the 10th session. This shows that in general although these aspects are achieved across the sessions, they were slow to become constant. This is related to the decline in the number of peer-imposed solutions, which, while not significant, do start form a trend towards the end of the study. This implies greater development of acceptance and tolerance and positive interdependence, given that an environment favoring equal communication among peers is essential for these to exist.

Motivation and interest is another variable that increased over time, eventually becoming constant after the fourth session (s.d. = 0.17 until the fourth session, versus. s.d. = 0.07 after the fourth session, p-value < 0.023). This indicates that it did not decrease as the sessions advanced, but instead reached a plateau. This also means it would not be affected even if the number of sessions were increased, suggesting that the students did not find the work to be tedious. In terms of the quality of brainstorming in

solving each exercise, this clearly increases and becomes constant following the 7th session. Although this increase is not significant ($p\text{-value} > 0.05$), it does imply that a greater number of sessions would not improve the quality of brainstorming achieved by the 7th session.

In terms of coordination, the students consistently required support from outside the group after the third session ($p\text{-value} < 0.027$). The level of disciplined work increased over time until the final session, although it never became constant. This could indicate that this particular aspect would continue to improve with further sessions. With regards to the students' appropriation of the platform, it is clear that their behavior towards the system became constant from the 4th session ($p\text{-value} < 0.00002$). This also indicates that the students' attitude towards the activity in general was positive and consistent. We can also observe that this is in direct relation to the achievements in motivation and interest, both of which remained constant from the 4th session onward. This also shows that working with the technological platform did not result in a loss of motivation towards the end of the study, despite the number of sessions. However, the suitable handling of material never became constant and instead increased continually until the end. As with the trend observed with disciplined work, this also suggests that the handling of materials could have improved with further sessions.

4. Discussion

Our first research question asked: when learning a language through integrated practice of the four skills, what advantages does a collaborative laboratory hold over an individual laboratory? As shown in the previous section of this paper, progress was made by all of the experimental groups (individual and collaborative language laboratory) in the four skills that were practiced: vocabulary, grammar, pronunciation, and listening. All of these results were significant, with the exception of the work on pronunciation in the individual laboratory. By comparing the results from the different groups (Table 5), we can see that pronunciation skills are always favored by the collaborative language laboratory, with a medium effect size (Cohen's $d > 0.62$). The use of collaborative learning in a language laboratory therefore contributed significantly to the development of pronunciation skills when compared to an individual language laboratory or a class without technology. These skills have been analyzed in the literature, particularly the use of computer mediation in developing listening and pronunciation skills (Bodnar, Penning de Vries, Cucchiarini, Strik, & Hout, 2011). Although the reported results show that these kinds of tools improve learning, they do not take into consideration the incorporation of collaborative work (Yamada, 2009).

The development of students' collaborative skills will be evaluated in the PISA test from 2015 onward. This will be done by measuring students' capacity and willingness to solve problems by interacting with each other (Davidson 2012; De Jong, 2012). The collaborative language laboratory's differentiating element is the opportunity it provides for peer-to-peer communication within the groups, based around solving the proposed exercises. This happens when the solution requires the students to reconstruct the sequence of a sentence, and where the semantics vary depending on the skill being practiced. Another opportunity for observed communication is when some of the peers do not pronounce the word correctly. The design of the activities enabled mutual reinforcement of pronunciation between participants based on shared listening of the recordings generated by each student in the group. This allowed students using the collaborative language laboratory to develop their pronunciation skills significantly better than the others students. However, the collaborative and individual language laboratories do not show improved results in the acquisition of grammar and vocabulary skills when compared to a lesson which does not use technology.

In response to the second research question, "when learning a language in a collaborative, face-to-face activity, which are the most relevant conditions of collaboration to be fulfilled?", we look to understand how collaboration assists learning. In order to do this, we analyzed how the conditions for collaboration detailed in the introduction are present in the proposed collaborative laboratory. This analysis was based on the results obtained in the in-class observations using the observation guidelines.

An important aspect to highlight is that in the 4th session, the values for the communication variables are higher than in the rest of the sessions. This could be due to the comparatively large number of words in the grammar activity, which required greater dialog to solve each task. This also affected several of the other qualitative variables such as mutual trust, acceptance and tolerance, motivation and interest, all of which produced lower values than in the other sessions. As mentioned in the previous section, this could be due to the influence of a greater number of peer-imposed solutions within the group.

In communication, Table 6 shows that positive interdependence is related to the larger number of person-to-person versus person-to-group dialogs. It is also related to the fact that each peer is responsible for the participation and learning of their classmates, based on the existence of a common goal. Another fact that supports this is that, while remaining constant over time, receiving help increases in proportion to the number of peer-imposed solutions, indicating that there was an awareness of the role played by peers within a group. This is also explained in terms of joint rewards, since the feedback is shared among peers and allows them to help one another.

As was mentioned in the results, the qualitative variables in the interaction, coordination and appropriation categories are closer to the maximum value, 3, than the

minimum value, 1. In interaction, we can observe how the constantly high value of the mutual trust variable across the sessions contrasts with the variables of acceptance and tolerance and positive interdependence. These latter variables do not become constant until the 10th session. This occurs because the students initially show behavior centered on accepting help from their peers, especially regarding the pronunciation of words. However, this becomes more participatory as there are fewer peer-imposed solutions and each student begins to acknowledge that their presence and the presence of others is essential to accomplish each task. This is confirmed by the peer-imposed solutions variable (communication) that decreases over time, with the exception of the 4th session, as explained previously. Although the decrease itself was not statistically significant, the decreasing value also allows us to confirm the existence of individual accountability. This is because the students participate more as they become more confident in sharing their work and ideas with the rest of the group in order to achieve the common goal. The development of individual accountability, evident in the decrease in the number of peer-imposed solutions, has been observed in cooperative language learning tasks where no single student can dominate proceedings and discourage other members from participating (AbuSeileek, 2012). This is also confirmed by the relation of the quality of brainstorming, acceptance and tolerance, and positive interdependence variables, which increase until the 7th and 10th sessions before becoming constant (p-value <0.0001, Table 6).

In the field of language learning, the use of positive interdependence and individual accountability have been analyzed separately, demonstrating that the latter presents significant advantages regarding the development of communication (AbuSeileek & AbuAlshar, 2012). In this study, the activity's design considers the use of mechanisms that attempt to fulfill all of the conditions for collaboration previously described. In a future study it would be interesting to analyze these conditions separately. This could be done by incorporating mechanisms that independently benefit each condition and using different versions of the same platform, as described by AbuSeileek & AbuAlshar (2012).

Furthermore, we conclude that better planning in solving the exercises (quality of brainstorming) shows acceptance and tolerance and positive interdependence. In the literature, the development of brainstorming has been considered as evidence of collaborative work and cooperation (Lee, 2011). It has also been considered as an opportunity for individual work prior to a face-to-face activity, where all students share their results (Long, 1990; Wen, Looi, & Chen, 2011). In the platform proposed in this project, the development of brainstorming is analyzed as a transversal element in the discussion generated by the students throughout the course of the activities. Based on this, we can suggest that the platform leads to higher quality discussions whenever the students plan how they would like to solve an exercise together. However, it is difficult to say which specific aspects could have benefited from this, since there is no phase explicitly designed for the development of brainstorming. This is something which as of yet has not been examined in the literature. Therefore, analyzing how specifically

incorporating the development of brainstorming among peers into the design of collaborative language learning activities remains an interesting topic for future work.

The improvement over time in disciplined work (coordination) allows the students to coordinate and communicate better among themselves, something which is also related to the gradual improvement in the quality of brainstorming. Requested support (requested mainly from the teacher) remained constant from the 3rd observed session onwards. This shows that their help was always necessary in solving the exercises, especially in the later sessions. Quality of brainstorming and requested support can be connected to studies that have evaluated learner autonomy in relation to their own learning process in on-line language learning. Here, the same trend is present over time (Dang & Robertson, 2010). Based on this, we can suggest that autonomy benefits from the platform proposed in this project. This is achieved by developing the quality of brainstorming and requests for support, which in turn is proof of coordination and communication between peers and individual accountability.

With regards to appropriation, the students were able to behave according to rules established by the activity almost from the beginning. This can be seen in the improvement in the behavior towards the system between the first and second observation (4th session) and the fact that it remains constant from then on. On the other hand, we observe that the gradual increase in the suitable handling of material variable also influenced the students' motivation and interest, the level of which increases over time (with the exception of the 4th session). This occurs because better use of the system helps the students to solve the exercises, and therefore improves their willingness to use the system as a team.

Finally, in terms of the specific work done with the orchestrations for the various activities, there was no rigorous follow-up. This was because this was not one of the main objectives of this study. However, the teacher in charge of leading the class did follow them and found them to be useful, making the necessary changes and modifying approximately a third of the original script. Most of these changes related to teacher interventions, and were made in order to allow the teacher to adapt the orchestrations to their own teaching style. This shows us that an orchestration is a valuable guide for the teacher, and that it should be sufficiently flexible for them to adapt it to their own needs. Future studies could be done on the impact of orchestrations on student learning and an analysis of to what degree this is accepted by the teachers.

In this study, activities were developed to develop vocabulary, grammar, and listening skills, with integrated pronunciation practice. The model used consisted of assigning objects to students, which then had to be put in order by following the defined logic (Zurita & Nussbaum 2004). The aforementioned skills can be practiced using other models, such as identifying, categorizing, or completing sequences or associations (Nussbaum, Rosas, Peirano, & Cardenas, 2001). It can be left as future work to enhance the laboratory using these various models and see which skills are better developed through collaboration in each case.

5. Conclusions

Our first research question was: when learning a language through integrated practice of the four skills, what advantages does a collaborative laboratory hold over an individual laboratory? We conclude that the proposed collaborative language laboratory can further improve English language learning, particularly pronunciation, when compared with an individual laboratory or a lesson without technology. Further studies must be carried out with a larger number of students, and spanning an entire school year in order to be able to evaluate the significance of these results. Our second research question was: when learning a language in a collaborative, face-to-face activity, which are the most relevant conditions of collaboration to be fulfilled? We conclude that among all of the potential aspects of a language learning activity, communication and coordination are the most important, since they form the basis of positive interdependence, the nucleus of good collaboration, and the mechanisms for accomplishing individual accountability and awareness.

References

- AbuSeileek, A. F. (2012). The effect of computer-assisted cooperative learning methods and group size on the EFL learners' achievement in communication skills. *Computers & Education*, 58(1), 231-239.
- Abuseileek, A. F., & AbuAlshar, A. M. (2012). The effect of computer-mediated communication cooperative learning structures and techniques on improving EFL learners' speaking skill. *International Journal of Learning Technology*, 7(4), 334-352.
- Adams, D. & Hamm, M. (1996). *Cooperative Learning: Critical thinking and collaboration across the curriculum* (2nd edition). Springfield, IL: Charles C. Thomas Publishers.
- Assessment and Teaching of 21st Century Skills (n.d.) Retrieved from <http://www.atc21s.org>
- Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, 211(4489), 1390.
- Barr, D., Leakey, J., & Ranchoux A. (2005). TOLD like it is! An evaluation of an integrated oral development pilot project. *Language Learning & Technology*, 9(3), 55-78.
- Bax, S. (2003). CALL - past, present and future. *System*, 31, 13-28.
- Blake, R. J. (2011). Current trends in online language learning. *Annual Review of Applied Linguistics*, 31(1), 19-35.
- Bodnar, S. E., Penning de Vries, B. W. F., Cucchiarini, C., Strik, H., & Hout, R. V. (2011). Feedback in an ASR-based CALL system for L2 syntax: A feasibility study. In *Speech and Language Technology in Education*.
- Berns, A., Palomo-Duarte, M., Dodero, J. M., & Valero-Franco, C. (2013). Using a 3D Online Game to Assess Students' Foreign Language Acquisition and Communicative Competence. In *Scaling up Learning for Sustained Impact*(pp. 19-31). Springer Berlin Heidelberg.
- Chen, C. M., & Chung, C. J. (2008). Personalized mobile English vocabulary learning system based on item response theory and learning memory cycle. *Computers & Education*, 51(2), 624 – 645.

- Chen, C. M., & Li, Y. L. (2010). Personalised context-aware ubiquitous learning system for supporting effective English vocabulary learning. *Interactive Learning Environments*, 18(4), 341-364.
- Chen, H. N. (2011). Developing and evaluating an oral skills training site supported by automatic speech recognition technology. *ReCALL*, 23, 1, 59–78.
- Chinnery, G. M. (2006). Going to the MALL: Mobile Assisted Language Learning. *Language Learning & Technology*, 10(1), 9-16.
- Chung, C. W., Lee, C. C., & Liu, C. C. (2012). Investigating face-to-face peer interaction patterns in a collaborative Web discovery task: the benefits of a shared display. *Journal of Computer Assisted Learning*.
- Dang, T. T., & Robertson, M. (2010). Impacts of learning management system on learner autonomy in EFL learning. *International Education Studies*, 3(3), P3.
- Davidson, M. (2012). *OECD International Assessment of Problem Skills*. Education for Innovative Societies.
- De Jong, J. H. (2012). Framework for PISA 2015: What 15-years-old should be able to do. *4th Annual Conference of Educational Research Center*. Broumana, Lebanon.
- Deutschmann, M., Panichi, L., & Molka-Danielsen, J. (2009). Designing oral participation in Second Life—a comparative study of two language proficiency courses. *ReCALL*, 21(2), 206-226.
- Dillenbourg P. (1999). What do you mean by collaborative learning?. In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and Computational Approaches* (pp.1-19). Oxford: Elsevier.
- Duncan, I., Miller, A., & Jiang, S. (2012). A taxonomy of virtual worlds usage in education. *British Journal of Educational Technology*, 43(6), 949-964.
- Garrido-Iñigo, P. & Rodriguez-Moreno, F. (2013). The reality of virtual worlds: pros and cons of their application to foreign language teaching. *Interactive Learning Environments*,. 1-18.
- Gutwin, C., & Greenberg, S. (2004). The importance of awareness for team cognition in distributed collaboration. In E. Salas & S. M. Fiore (Eds.), *Team cognition: Understanding the factors that drive processes and performance* (pp. 177–201).
- Hampel, R. (2009) Training teacher for the multimedia age: developing teacher expertise to enhance online learner interaction and collaboration. *Innovation in language learning and teaching*, 3(1), 35-50.

- Handley, Z. (2009). Is text-to-speech ready for use in computer-assisted language learning? *Speech Communication*, 51, 906-919.
- Harvey, T. E. (1978). The matter with listening comprehension isn't the ear: hardware & software. *NALLD Journal*, 13(1), 8-16.
- Hemard, D. (1997). Design principles and guidelines for authoring hypermedia language applications. *System*, 25 (1), 9-27.
- Hsu, C.K., Hsu, C.Y., He, Y.Y. & Chang, C.K. (2009). Verification of the Perception and Optimal Number of Learners in a Group for Computer-Assisted Language Learning. In G. Siemens & C. Fulford (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009* (pp. 1720-1735). Chesapeake, VA: AACE.
- Hsu, S.Y. (2005). Building Language-Learning Environments to Help Technological University Students Develop English Independent Learning. *The JALT CALL Journal*, 1 (2), 51-66.
- Infante, C., Hidalgo, P., Nussbaum, M., Alarcón, R. & Gottlieb, A. (2009). Multiple Mouse Based Collaborative One-to-One Learning. *Computers and Education*, 53(2), 393-401.
- Janssen, J., Erkens, G., Kanselaar, G., & Jaspers, J. (2007). Visualization of participation: Does it contribute to successful computer-supported collaborative learning? *Computers & Education*, 49(4), 1037-1065.
- Johnson, D. W., & Johnson, R. T. (1999). Making cooperative learning work. *Theory Into Practice*, 38(2), 67-73.
- Kaplan, F., DoLenh, S., Bachour, K., Yi-ing Kao, G., Gault, C. & Dillenbourg, P. (2009) Interpersonal Computers for Higher Education. In P. Dillenbourg et al. (Eds.), *Interactive Artifacts and Furniture Supporting Collaborative Work and Learning*, 10, 1-17, Springer.
- Kessler, G., & Bikowski, D. (2010). Developing collaborative autonomous learning abilities in computer mediated language learning: attention to meaning among students in wiki space. *Computer Assisted Language Learning*, 23(1), 41-58.
- Leahy, C. (2008). Learner activities in a collaborative CALL task. *Computer Assisted Language Learning*, 21(3), 253-268.
- Leahy, C. (2012). Learner Behavior in a Collaborative Task-Based CALL Activity. In Call: Using, Learning, Knowing, EUROCALL Conference, Gothenburg, Sweden, 22-25 August 2012, Proceedings (p. 167). Research-publishing. net.

- Lee, K.W. (2000). English teachers' barriers to the use of Computer assisted language learning. *The Internet TESL journal*, 6(12). Retrieved from <http://iteslj.org/Articles/Lee-CALLbarriers.html>
- Lee, L. (2011). Blogging: Promoting Learner Autonomy and Intercultural Competence Through Study Abroad. *Language Learning & Technology*, 15(3), 87-109.
- Levy, M. (2009). Technologies in Use for Second Language Learning. *The Modern Language Journal*, 93(1), 769-782.
- Li, L., & Wong, D. (2010). What Can Avatars Do? Virtual realities in collaborative language learning. In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2010* (pp. 466-474). Chesapeake, VA: AACE
- Liaw, S. S. (2002). An internet survey for perceptions of computer and world wide web: relationship, prediction, and difference. *Computers in Human Behavior*, 18(1), 17-35.
- Liaw, S. S., Chen, G. D., & Huang, H. M. (2008). Users' attitudes toward Web-based collaborative learning systems for knowledge management. *Computers & Education*, 50(3), 950-961.
- Long, M. (1990). Task, group, and task-group interactions. In S. Anivan (Ed.), *Language teaching methodology for the nineties*. Singapore: SEAMEO Regional Language Centre
- Mackey, A. (2006) Feedback, Noticing and Instructed Second Language Learning. *Applied Linguistics*, 27(3), 405–430.
- Merisuo- Storm, T. (2007). Pupil's attitude toward foreign-language learning and the development of literacy skills in bilingual education. *Teaching and Teacher Education*, 23(2), 226-235.
- Mineduc (2010). Bases curriculares Ingles 6to básico. Retrieved from <http://www.curriculumenlineamineduc.cl/605/w3-propertyvalue-52067.htm>
- Morton, F. (1960). The Language Laboratory as a Teaching Machine. *Annual Bulletin 19 of the Connecticut Audio-Visual Education Association*, "Modern Techniques in Teaching Foreign Languages: Language Laboratories", 162 – 165.
- Neri, A., Cucchiarini, C., & Strik W. (2003). Automatic speech recognition for second language learning: How and why is actually works. In *Proceedings of the 15th international Conference on Phonetic Sciences* (p. 1157 – 1160), Barcelona:[s.n].
- Nguyen, L. V (2010). Computer Mediated Collaborative Learning within a Communicative Language Teaching Approach: A Sociocultural Perspective. *Asian EFL Journal*, 12(1), 202 – 233.

- Nussbaum, M., Dillenbourg, P., Dimitriadis Y., and Roschelle, J., (2013) Guest Editors: Classroom Orchestration, *Computers & Education*, Volume 69, pp 485- 523
- Nussbaum, M., Rosas, R., Peirano, I., Cardenas, F. (2001). Development of intelligent tutoring systems using knowledge structures. *Computers & Education*, 36, 15-32
- Ogata, H. (2008). Computer Supported Ubiquitous Learning: Augmenting Learning Experiences in the Real World. In *Fifth IEEE International Conference on Wireless, Mobile, and Ubiquitous Technology in Education WMUTE 2008* (pp. 3-10).
- Ogata, H., & Yano, Y. (2003). How Ubiquitous Computing can Support Language Learning. In *Proceedings of KEST 2003* (pp.1-6).
- Peterson, M. (2012). EFL learner collaborative interaction in Second Life, *ReCALL*, 24(1), 20-39.
- Plass, J. & Jones, L. (2005). Multimedia Learning in Second Language Acquisition. In R. Mayer (Ed), *The Cambridge handbook of multimedia learning* (pp. 467-488). Cambridge: Cambridge University Press.
- Pranita, G. (2010). Digital Language Laboratory and the Teaching of English in India. *Educational Quest- An International Journal of Education and Applied Social Sciences*, 1(1), 39-45.
- Rama, P. (2012). Review of “Language Learning in the Digital Age”. *Language Learning and Technology*, 16(1), 30-33.
- Roby, W. B. (2004). Technology in the Service of Foreign Language Learning: The Case of the Language Laboratory. In David Jonassen (Ed.): *Handbook on Research on Educational Communications and Technology* (pp. 523-542). Mahwah, NJ: Lawrence Erlbaum Associates.
- Saggara, N., Zapata, G. C. (2008). Blending classroom instruction with online homework: A study of student perceptions of computer- assisted L2 learning. *ReCALL*, 20(2), 208-224.
- Shih, Y. C., & Yang, M. T. (2008). A collaborative virtual environment for situated language learning using VEC3D. *Educational Technology & Society*, 11(1), 56–68.
- Singhal, M. (1997). The Internet and Foreign Language Education: Benefits and Challenges. *The Internet TESL Journal*, 3(6). Retrieved from <http://iteslj.org/Articles/Singhal-Internet.html>.
- Sykes, J. (2005). Synchronous CMC and pragmatic development: Effects of oral and written chat. *CALICO Journal*, 22, 399–431.
- Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology*, 21(1), 43–69.

- Stewart, J., Bederson, B., & Druin, A. (1999). Single Display Groupware: A Model for Co-present Collaboration. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 286-293).
- Szewkis, E., Nussbaum, M., Rosen, T., Abalos, J. Denardin, F., Caballero, D., Tagle, A. & Alcoholado, C. (2011) Collaboration between large groups in the classroom. *International Journal of Collaborative-Supported Collaborative Learning*, 6(4), 561-575.
- Vanderplank, R. (2009). Déjà vu? A decade of research on language laboratories, television and video in language learning. *Language Teaching*, 43(1), 1-37.
- Vandewaetere, M. & Desmet, P. (2009). Introducing psychometrical validation of questionnaires in CALL research: the case of measuring attitude towards CALL. *Computer Assisted Language Learning*, 22(4), 349-380.
- Wagener, D. (2006). Promoting independent learning skills using video on digital language laboratories. *Computer Assisted Language Learning*, 19(4 & 5), 279–286.
- Wang, Y., & Chen, N.-S. (2010). The collaborative language learning attributes of cyber face-to-face interaction: the perspectives of the learner. *Interactive Learning Environments*, 1-20.
- Warschauer, M., & Healey, D. (1998). Computers and language learning: an overview. *Language Teaching*, 31, 57 – 71.
- Warschauer, M. (1997). Computer-mediated collaborative learning: theory and practice. *The Modern Language Journal*, 81(4), 470-481.
- Wen, Y., Looi, C. K., & Chen, W. (2011). Towards a model for rapid collaborative knowledge improvement in classroom language learning. *Connecting Research to Policy and Practice: Proceedings of CSCL*, 836-851.
- Wright, B. D., & Stone, M. H. (1979). Best Test Design. Rasch Measurement.
- Xu, Y., & Seneff, S. (2009). Speech-based interactive games for language learning: reading, translation, and question-answering. *Computational Linguistics and Chinese Language Processing*, 14(2), 133-160.
- Yamada, M. (2009). The role of social presence in learner-centered communicative language learning using synchronous computer-mediated communication: Experimental study. *Computers & Education*, 52(4), 820-833.
- Yang, Y.-T. C., Gamble, J., & Tang, S.-Y. S. (2011). Voice over instant messaging as a tool for enhancing the oral proficiency and motivation of English-as-a-foreign-language learners. *British Journal of Educational Technology*. doi: 10.1111/j.1467-8535.2011.01204.x.

- Zheng, D., Young, M. F., Wagner, M., & Brewer, R. A. (2009). Negotiation for Action: English Language Learning in Game-Based Virtual Worlds. *The Modern Language Journal*, 93(4), 489-511.
- Zurita, G., & Nussbaum, M. (2004a) A Constructivist Mobile Learning Environment Supported by a Wireless Handheld Network. *Journal of Computer Assisted Learning*, 20, 235-243.
- Zurita, G., & Nussbaum, M. (2004b) MCSCL: Mobile Computer Supported Collaborative Learning. *Computers & Education*, 42 (3), 289-314.

Appendix 1: Example of orchestration

Class 2 Orchestration

Subject Area: Foreign Language: English	Unit Name: My Family	Designated Time: 1 lesson (45 minutes)
---	--------------------------------	---

Number of Lessons Per Week:	Grade (s): 6
------------------------------------	---------------------

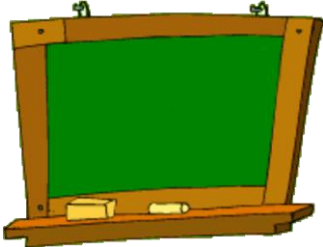
Difficulty: (X) Easy () Medium () Difficult	Language Skills: (X) Listening () Reading (X) Speaking	
---	---	--

Unit Learning outcomes: Demonstrate understanding of commonly used phrases and expressions in brief and simple dialogs or oral presentations using direct language.

Skills and Attitudes to develop:

Listening comprehension.			
Speaking			
Class Objective: Present family members vocabulary, then use it to perform reading comprehension and oral production tasks.			
Necessary Prior Learning Skills: Personal Pronouns: I, he, she, it, we, you, they. Vocabulary: Members of the (nuclear) family			
Class Stage	Class Activities	Teaching Materials	Designated Time

Beginning of Class: Introduction of Contents	 The teacher begins the class showing students a family photo where they can identify the following members: father, mother, brothers and sisters, grandparents, among others. Then the teacher asks: - Who are the people in this photograph? How are they related to me? Students are expected to respond orally to these questions, making suggestions about the picture and the people that appear in it. Then students are expected to identify that the picture is of the teacher's family and identify the names of the family members. Next, the teacher explains to students what today's class is about: - Today we are going to review vocabulary about family members in an activity on the computer to practice and use what we've learned.	-Family Photo	5 minutes
--	--	----------------------	-------------------------

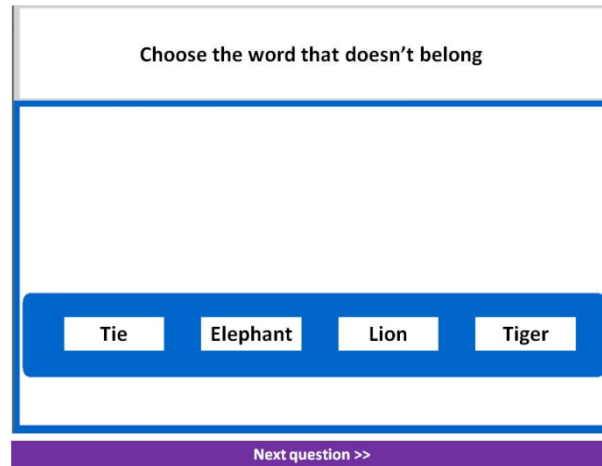
Class Stage	Class Activities	Teaching Materials	Designated Time
Class Progress: Exercising Contents	<i>This part of class consists of 2 activities: brainstorming and recall on the board, and a listening exercise integrating the unit vocabulary, with supplementary ICT resources.</i> Activity 1: The teacher draws a tree on the whiteboard and asks students if they know what a family tree is and how to make one.  The teacher invites the class to make a tree together, writing the names of family members in Spanish. The objective here is to recall keywords referred to in the unit. Next, the teacher sticks a label with the name of each family member in English next to each name in Spanish. Then the students are invited to practice pronunciation, repeating the words individually and as a group. Once the students have understood and practiced pronouncing the vocabulary, the teacher must give the following instruction: <i>-Now that we've remembered the keywords for this topic, we're going to work in groups on the computer practicing this vocabulary by listening to a story. Let's go to the computer room.</i>	- Whiteboard. - Marker. -Labels (names of family members in English)	Activity 1.5 minutes

Class Stage	Class Activities	Teaching Materials	Designated Time
Class Progress: Exercising Content	Activity 2: Next the teacher invites the students to work in groups, listening to a paragraph and ordering three words that appear in it. The teacher reminds students that listening closely, identifying words, their order, coming to a consensus and practicing pronunciation are crucial in the activity.  During the course of the activity the teacher: <u>Monitors</u> the <u>students' work and helps</u> those that appear to need more help with pronunciation. Registers the words with the highest level of appropriation in pronunciation, to include in future activities. Registers the phonemes that cause the greatest problems in pronunciation, to work on in future activities.	Computer room set up for the task. -Listening exercises: 3, 7, 14, 16, 21,22,24	Activity 2.3 minutes

Class Stage	Class Activities	Teaching Materials	Designated Time
End of Class: Final Content Review	<i>At this time the class consists of 1 activity: oral expression in a group setting and summary of the contents of the class.</i> Activity 1: The teacher <u>invites</u> students to describe a member of their family physically, in English. In order to do this, the teacher writes a list of physical features and the following phrase on the board: “My...is...” Students are expected to mention a member of their family and the feature, based on the work done in class. At this time the teacher monitors and registers the appropriation of vocabulary practiced in class, and makes a list of the most repeated words to start introducing new words in the following classes.	-No extra materials required.	Activity 1: 5 minutes

Figure A.1: “My family” orchestration

Appendix 2: Example of pre-post test items



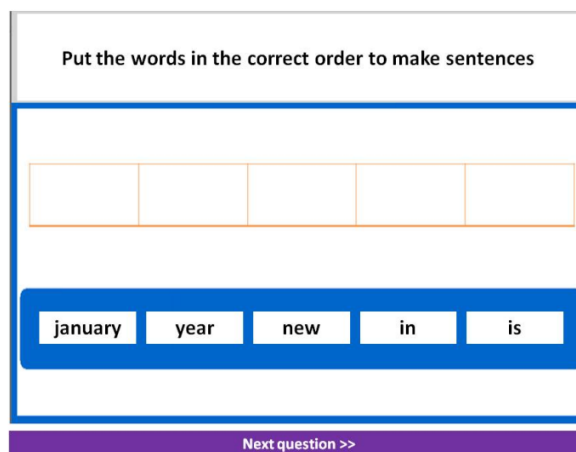
Choose the word that doesn't belong

Tie Elephant Lion Tiger

Next question >>

The interface for a vocabulary test item. It features a light gray header with the instruction "Choose the word that doesn't belong". Below this is a large white rectangular area for the user's selection. At the bottom of this area is a blue horizontal bar containing four white buttons labeled "Tie", "Elephant", "Lion", and "Tiger". A purple footer bar at the very bottom contains the text "Next question >>" in white.

Figure B.1: Vocabulary item



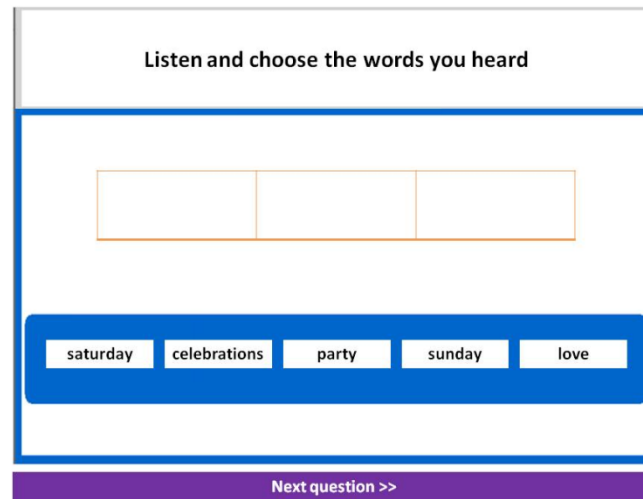
Put the words in the correct order to make sentences

January year new in is

Next question >>

The interface for a grammar test item. It features a light gray header with the instruction "Put the words in the correct order to make sentences". Below this is a large white rectangular area containing five empty orange-outlined boxes for the user to place the words. At the bottom of this area is a blue horizontal bar containing five white buttons labeled "January", "year", "new", "in", and "is". A purple footer bar at the very bottom contains the text "Next question >>" in white.

Figure B.2: Grammar item



Listen and choose the words you heard

Three empty rectangular boxes for writing answers.

saturday celebrations party sunday love

Next question >>

The interface consists of a white header box with the text "Listen and choose the words you heard". Below this is a large white area containing three empty rectangular boxes for writing answers. At the bottom of this area is a blue rounded rectangle containing five white buttons with the words "saturday", "celebrations", "party", "sunday", and "love". The entire interface is framed by a blue border, and a purple footer bar at the bottom contains the text "Next question >>".

Figure B.3: Listening item



Say the word

Socks

Next question >>

The interface consists of a white header box with the text "Say the word". Below this is a large white area. At the bottom of this area is a blue rounded rectangle containing a white button with the word "Socks". The entire interface is framed by a blue border, and a purple footer bar at the bottom contains the text "Next question >>".

Figure B.4: Pronunciation item

Appendix 3: Observation guideline

Observation guideline

		STUDENTS	GROUP:	
Session/ Date:			1.	
Level:	6 th Grade		2.	
Assignment:	English		3.	
Unit:				

Start **End**

Category	Variable	Metric	Type*	1	2	3 / G
Communication	Person to person	# dialogs	I			
	Person to group	# dialogs	I			
	Receive support	# requests	I			
	Request for support	# requests	I			
	Peer-imposed solution	# events	G			
Interaction	Positive interdependence	Scale of 1 to 3	G			
	Mutual trust	Scale of 1 to 3	G			
	Acceptance and tolerance	Scale of 1 to 3	G			
	Motivation and interest	Scale of 1 to 3	G			
	Quality of brainstorming	Scale of 1 to 3	G			
Coordination	Disciplined work	Scale of 1 to 3	G			
	Requested support	Scale of 1 to 3	G			
Appropriation	Suitable handling of material	Scale of 1 to 3	G			
	Behavior towards the system	Scale of 1 to 3	G			

* (I) Individual; (G) group work

Figure C.1: Observation guideline.

**ANEXO B: ARTÍCULO ENVIADO A LA REVISTA JOURNAL OF RESEARCH
IN READING**

Journal of Research in Reading



**Comic-Based Dynamic Assessment Instrument for
Measuring Reading Comprehension among Third Grade
Students**

Journal:	<i>Journal of Research in Reading</i>
Manuscript ID	JRIR-2016-03-0053
Manuscript Type:	Original Article
Keywords:	Assessment, Reading comprehension, comic

SCHOLARONE™
Manuscripts

Comic-Based Dynamic Assessment Instrument for
Measuring Reading Comprehension among Third Grade Students

Keywords: reading comprehension, dynamic assessment, comic, interactive learning

Abstract

Assessing reading comprehension requires multiple indicators in order to measure underlying variables and avoid issues with construct validity. Dynamic assessment methods may provide a good alternative as they are capable of adapting to and measuring a student's level. In addition to this, students are drawn to comics through intrinsic motivation, making them attractive as an assessment instrument. The research question that guides this study therefore asks "How effective is a comic-based, dynamic assessment instrument at measuring reading comprehension?" To answer this, an instrument was developed for use with tablets and validated internally and externally using a standardized test and the students' grades. The dynamic assessment instrument correctly measured the students' level of reading comprehension, providing the results in real time. Using this type of instrument in the classroom facilitates Formative Assessment, using the students' actions and mistakes as an opportunity for learning, as well as to guide the teacher's activities.

Keywords: reading comprehension, dynamic assessment, comic, interactive learning

Highlights

What is already known about this topic:

- Assessing reading comprehension requires multiple indicators in order to measure underlying variables and avoid issues with construct validity.
- The methods for assessing reading comprehension have not undergone any great changes in recent years. Assessment is mainly based on standardized pen-and-paper tests.
- Dynamic assessment methods are capable of adapting to each student and their respective level of knowledge.

What this paper adds:

- A student is drawn to a comic through intrinsic motivation, especially if it is adapted with their interests in mind, making them attractive as an assessment instrument.
- When compared to a standardized test, a comic-based dynamic assessment instrument has a similar capacity to assess a student's level of reading comprehension.

Implications:

- Methods for assessing reading comprehension could evolve not only into the digital world, but also take advantage of new resources and implement dynamic assessment to encourage student development.
- The type of test proposed in this research allows the student's level to be measured and understood, providing a base on which to improve. Furthermore, the instrument provides results in real time and can be accessed by a large number of students at the same time through the internet, making it accessible and reducing the costs of distribution and application.

Comic-Based Dynamic Assessment Instrument for Measuring Reading Comprehension among Third Grade Students

Reading comprehension is considered an essential skill; developing this skill is one way for people to acquire important tools for their professional, academic and social life. Chile is the top-scoring country in Latin America and the Caribbean for reading comprehension and mathematics in third grade (TERCE, The Latin American Laboratory for Assessment of the Quality of Education, 2015). However, 84% of the population do not understand what they read and 30% of Chilean students do not acquire the basic reading skills that are needed to play a full role in society, either now or in the future (Center for Microdata, Economics Department, Universidad de Chile, 2011). The United States Department of Education suggest that if a child is not able to read fluently by 4th grade, there is approximately a 78% chance that they will not catch up for the remainder of their education (National Center for Education Statistics, U.S. Department of Education, 2007).

There are several levels of knowledge and performance involved in reading and that are acquired through practice. A reader starts with micro-skills, such as identifying the smallest units of a language, e.g. graphemes (letters) and phonemes (sounds), before advancing to chains of morphemes (parts of words), semantics (meaning) and syntax (grammatical structures), until finally managing to understand sociocultural contexts (types of discourse). The latter are referred to as macro-skills (Brown, 2004). It is necessary to measure both macro and micro skills in order to identify a person's overall level of effective reading. Another model, based on so-called "pathways", has also been put forward. In this case, the "pathways" reference the three areas of learning that must

be incorporated when teaching early reading skills: decoding, building meaning; and knowledge of genres and types of text, from a sociocultural perspective (McKenna & Stahl, 2012).

Reading comprehension is a multi-faceted construct. The choice of text and the tasks that are included in an assessment will, to a certain extent, drive the results that are obtained (O'Reilly et al., 2014). Multiple indicators must be included when assessing reading so as to accurately measure the underlying variables (Fletcher, 2006). If this does not happen, the results of the assessment may experience issues with bias or construct validity (Cook & Campbell, 1979). The methods for assessing reading comprehension have not undergone any great changes in recent years. Assessment is mainly based on standardized pen-and-paper tests. Given the static nature of such tests, they are only capable of identifying the students' existing skills and do not take into consideration the differences among students (Nazari & Mansouri, 2014). This latter point is particularly relevant when considering that, nowadays, the international literature reveals that effective early reading practices take into account the students' individual differences (Connor, Morrison & Katch, 2004; Snow, Griffin & Burns, 2005; Connor et al., 2009).

In contrast, dynamic assessment methods are capable of adapting to each student and their respective level of knowledge. Such methods can even determine a student's learning potential, which Vygotsky refers to as the zone of proximal development (ZPD) (Vygotsky, 1978). In addition to providing an assessment, these dynamic methods also promote the development of reading comprehension skills among students, through a clinical method, whereby the mediator and the student collaborate in order to conduct an

assessment (Nazari & Mansouri, 2014). Static measurements only reveal what has already been learned or achieved, while the range of the ZPD provides possible indicators of what may be learned in the future. Dynamic assessments therefore provide the possibility to measure and understand a student's level and, based on this, to improve their level of development. For a comparison between static and dynamic methods, see Table 1 in (Nazari & Mansouri, 2014).

An interest in new types of literacy has developed in recent decades. This has been driven by the opportunities provided by new technology, which serves as an aid to display hybrid or multimodal texts (Cope & Kalantzis, 2009). As a result of this digitalization, visual modes such as image and film, along with other such expressive modes, have been increasingly valued and incorporated into teaching. Furthermore, in some cases, these alternative modes can now be seen as being favoured over the more traditional verbal/written mode (Kress, 2003). We therefore see the need to teach literacy beyond the confines of oral and written language; extending teaching to other modes, such as aural, tactile, gestural and spatial. School-age children of today are sensitive to a multimodal world, though schools do not always foster their sensorial skills in order to understand this type of discourse. In such situations, these skills could be used as a base from which to work towards multimodal comprehension (Cope & Kalantzis, 2009; Gee, 2004).

Given how students typically react to a comic, they would appear to be a potentially attractive instrument for measuring reading comprehension. Including visual graphics and images along with a text not only helps motivate the student; it can also help develop visual literacy (Valerie & Abed, 2013). A student is drawn to a comic

through intrinsic motivation, especially if it is adapted with their interests in mind (Cimermanová, 2015). Comics boast the following characteristics, making them particularly valid as a measurement instrument (Cuadrado et al., 1999): they are easier to understand than verbal texts, they appeal to both teachers and students, they are fun, they excite students, they have an impact, they encourage creativity, they favour a classroom dynamic, they can be taken in at a glance, and they add colour to a class. In addition to this, the multimodal nature of the comic is well suited to digital media, which can stimulate a student's interest in reading (Frey & Fisher, 2004). Given the above, the research question that guided this study asks “How effective is a comic-based, dynamic assessment instrument at measuring reading comprehension?”

Methodology

Instrument design: Using comics as a multimodal and dynamic assessment instrument to measure reading comprehension

A digital, dynamic instrument was designed and developed for tablets. The aim of this instrument was to measure the level of reading comprehension among primary school students. This instrument was based on a comic, with a multimodal narrative that was an adaptation of one of the texts recommended by the Chilean Ministry of Education for students in 1st to 6th grade: “Around the World in 80 Days” by Jules Verne (MINEDUC, 2012). “Around the World in 80 Days” originally contains 42 chapters, which were condensed into 17 sections, developed as comic strips that followed the story's natural sequence. As a multimodal genre, comics are made up of fixed panels

that are preferably horizontal, although they can also be vertical. The characters can also be fixed, and predominantly use dialogue as part of a sequenced narrative (Evans, 2013).

The skills required to understand written discourse are multi-componential, working on several, complementary levels of processing: syntactic, lexical-semantic and discourse, such as the type of text and/or genre (Koda, 2005). Given this, in order to design and create the comic, a series of grids containing elements from the story (narrative categories) were created, as well as textual and lexical/grammatical grids. These were then cross-referenced with three levels of reading difficulty: basic, intermediate and advanced, in order to develop questions for each level.

In more specific terms, this meant first developing a grid of elements on a macro level, in the form of a “story grammar” or “story structure”. In order to do so, the elements were assigned basic narrative categories, such as setting, characters, theme and plot. Subsequently, the story was written for the three aforementioned difficulty levels, based on a macro-level (sentence/phrase) grid of elements and an appropriate level of grammatical and lexical complexity. This allowed three simultaneous stories to be developed, each of a different length and level of complexity. The vocabulary used was taken from the word lists suggested by the Ministry of Education for third grade. Furthermore, these grids were then used to create an assessment rubric, which allowed questions to be developed for each difficulty level. Doing so therefore ensured that the difficulty level was only determined by comprehension of the text and not by the syntactic/lexical complexity of the questions. Each difficulty level is associated with its own type of questions: literal or textual questions for the basic level, simple inference questions for the intermediate level, and general or complex inference questions for the

advanced level. Literal or textual questions require students to locate explicit information in the text or images, or related to the text/images, and which may even be highlighted in some way within the text. Simple inference questions require an understanding of the meaning of the phrases and words by using the context, establishing cause-effect relationships, making comparisons and contrasts, among other processes. Complex inference questions are often focused on making inferences based on textual clues, main and secondary ideas, as well as predicting actions, conclusions or outcomes (León & Pérez, 2003). These levels match with those that are assessed on a macro level in the majority of international standardized (and validated) tests, such as PISA and PIRLS, as well as the local SEPA test (MIDE-UC, 2009a, 2009b) that was used in this study. These levels are often specified in more detail, resulting in descriptors that are grouped into as many as six levels, such is the case with PIRLS.

Therefore, according to the student's level, the questions vary in their level of difficulty. Table 1 shows a sample of the questions for each section. Once the student finishes a section, a character asks them a question and provides them with three alternatives, of which they must choose one. Figure 1 shows an example of the question box.

When the student correctly answers a question, they advance to the next section, where the text and questions are more difficult, unless the student is at the advanced level, in which case the difficulty level remains the same. When the student makes a mistake they must read the same section again, but this time the text and questions are easier. Should the student make two consecutive mistakes, the system lowers the difficulty level again, unless the student is at the basic level, in which case the difficulty

level remains the same. The flow between the different difficulty levels is detailed in Figures 2 and 3. The length of the story that is finally read by the student depends directly on their level of reading comprehension and the amount of mistakes they make during the activity.

The assessment session begins with the student identifying themselves by selecting their name from a pre-configured list. The system stores their progress as they work their way through the comic. This allows for interruptions in the activity as the student can leave the program and then continue from where they left off. It also allows the student to advance at their own pace and according to their specific needs.

The instrument's display is organized into 3 areas on the screen (see Figure 4):

1. Previous panel (see Figure 4, area "A"): A space which partially reveals the contents of the last panel that was read. The aim of this is to provide the look and feel of a comic strip. At the start of each section this area appears blank (see Figure 5, area "A").
2. Main contents (see Figure 4, area "B"): A space which shows what the student is currently reading, i.e. the latest panel. This space changes every time the student gets to the end of a section, in order to present the student with the reading comprehension question. The question appears in the top left corner of this area (see Figure 1, area "A") with the 3 possible answers appearing in the top right corner (see Figure 1, area "B").
3. Next panel (see Figure 4, area "C"): A space which partially reveals the contents of the next panel that will be read by the student. The aim of this is

to provide the look and feel of a comic strip. At the end of each section this area appears blank (see Figure 1, area “C”).

During and at the end of the reading activity, the teacher using the instrument can access a progress report for each student (see Figure 6). This report shows the percentage of correct answers given by each student for the respective levels, the time spent on the assessment and whether the answer given to each question is correct or incorrect.

This report shows the difficulty level of each question answered by the student, providing the teacher with a graph of the student’s performance throughout the session and therefore their level of reading comprehension. In Figure 6 we can see that the student Ignacio Carmach correctly answered the questions and stayed at the advanced level until section 9, where he made a mistake and therefore repeated the section, though at an intermediate level and with questions aimed at this level. In the example, the student correctly answered the question and continued to the next section. In this section, the student again answered incorrectly and therefore dropped down to a basic reading level. This situation is repeated in section 11, but as the student is already at the basic level, the responses are shown in a single cell on the report: first a mistake and then a correct answer. The student then continues through the story, answering the questions correctly and working his way up to the advanced level until section 16, where he makes his final mistake before finishing the story and ending the activity at the intermediate level.

Sample

The instrument was tested in two state schools and one private school in Santiago, Chile, with 198 third grade students (90 girls, 108 boys), aged between 9 and 10.

The schools were chosen from among those that administer the standardized SEPA test for Language Arts (MIDE-UC, 2009c). This test is aligned with Chile's national curriculum and provides information regarding student achievement and progress throughout the school year (Alfonso et al., 2010). This selection was made with the aim of working with schools that assess their students using a standardized test. By doing so, it was possible to measure the effectiveness of the instrument that was developed for this study, comparing the results in order to establish the instrument's construct validity. Given the differences between the static SEPA test and the dynamic test developed for this study, a comparison was made between the achievement levels measured by the static test, both in the various areas assessed by the test as well as their overall average, and the results produced by our instrument. In addition to the results from the SEPA test, the students' grades in Language Arts were also used in order to validate the instrument.

Technology platform used

10.1" tablets were used for this study. Anxiety regarding the use of this technology was only observed in approximately 20% of the students in each class. In general terms, the students showed great enthusiasm for and interest in the activity.

Procedure

Before the study began, the SEPA test was applied independently by the teachers in each school, who followed a protocol that was provided by the SEPA team. The test lasted 90 minutes and was taken a few months (between 1 and 2 months) before the instrument was used at each school.

The students' level of reading comprehension was assessed in each class using the instrument that was developed for this study. The instrument took approximately 45 minutes to administer in each class, with students taking between 13 and 45 minutes to complete the activity.

The results from the SEPA test, the report on each student generated by the system used in this study and the students' grades in Language Arts (corresponding to the period in which the instrument was used in each school), were used to validate.

The tables with the results from the SEPA test, the results from the instrument and the students' grades are not included in this paper. These can be found at the following website: <https://goo.gl/D8rYp0>.

Analysis plan

Descriptive analysis was performed on the achievement levels for both tests. Cronbach's alpha was used to establish the instrument's internal consistency by considering each section as an item. The item-test correlation was then calculated in order to determine whether any of the items needed to be eliminated. Simple linear regression was used to establish the instrument's external validity, using two types of data at a time. In order to establish favourable correlations, the results obtained by the students on the instrument were compared with the percentage scores on the SEPA test, and then separately with the grades in Language Arts. Finally, the results obtained by the

students on the instrument were analysed using a two-way ANOVA in order to establish whether there were any significant differences between boys and girls and thus any gender bias.

Results

Analysis of the instrument's internal validity

Cronbach's alpha for the instrument is 0.829. Table 2 reveals the analysis of the instrument's reliability, which takes into consideration removing each of the sections. This analysis suggests that the reliability does not improve if any of the sections is removed. All of the sections are therefore kept in order to ensure that the story progresses smoothly.

Analysis of the instrument's external validity

Using simple linear regression (see Table 3), positive and favourable correlations were found between the results obtained on the instrument developed for this study and the students' grades in Language Arts, explaining 24.1% (Adjusted R squared) of the variables that determine said performance. A positive correlation was also found between the results obtained on the instrument and the students' percentage scores on the SEPA test, this time explaining 31.5% of the variables that determine said performance.

In this sense, in both cases such a percentage is not insignificant given that academic performance is determined by a large number of variables from different levels (individual student level, school level, community level, etc.) (Taut et al., 2014).

An ANOVA test (see Table 4) was used to determine whether the instrument had any gender bias. The results of this test revealed significant differences between the average scores by boys and girls for each of the assessments. In other words, this was the case for the dynamic assessment system ($P = 0.0305$), the students grades in Language Arts ($P = 0.00103$), and the percentage scores on the SEPA test ($P = 0.0114$). These results are consistent with those reported for the national Chilean standardized test for Language Arts in 4th grade (SIMCE), which also reveals differences in favour of girls.

Discussion

Based on the above results, it can be concluded that the instrument designed for this study correctly assessed the level of reading comprehension for all of the students in the sample using macro-skills. This was true regardless of the student's starting level. As the instrument allowed the students to advance at their own pace, a significant gap was produced between the students who started at an advanced level and those that started at a basic level, which was reflected in the time it took each student to complete the test (see Table 5). Despite this, it is worth noting that the level of reading comprehension was correctly measured in every case.

During the experimental sessions, the students showed signs of tiredness and lack of concentration after 20 minutes. Considering that it took the students an average of 25 minutes to complete the test, this may explain why the students made more mistakes in the final sections. Given this, for future applications of such instruments it is recommended limiting the length of the story.

Each session was conducted by the researchers. However, each teacher could administer the test as the instrument is accessible through a website without the need for any additional configuration.

Conclusions

In this study we asked “How effective is a comic-based, dynamic assessment instrument at measuring reading comprehension?” Through triangulation with data from different sources, the results reveal the instrument’s predictive capacity. Given these results, the dynamic instrument and the standardized, static test both have a similar capacity to assess a student’s level of reading comprehension. However, the dynamic instrument encourages student development (Nazari & Mansouri, 2014) and allows the student’s level to be measured and understood, providing a base on which to improve. Furthermore, the instrument provides results in real time and can be accessed by a large number of students at the same time through the internet. Finally, as the questions are contained in a comic strip and thanks to the connection that this produces with the reader (Cuadrado et al., 1999), the instrument should lower the levels of nervousness that standardized tests cause on students. This would in turn lead to more reliable assessments as there would be less variability. It remains as future work to test this hypothesis.

Given the way in which the sample was chosen, the number of schools and the range of subjects analysed, the results of this study cannot be generalized. In order to generalize the results, a cross-curriculum study involving a larger, randomly-selected sample of students from several grade levels is required. Finally, using this type of

instrument in the classroom facilitates Formative Assessment, using the students' actions and mistakes as an opportunity for learning, as well as to guide and modify the teacher's activities in the classroom (Black & William, 1998). It remains as future work to study how an instrument such as the one in this study can be used to teach a subject through formative assessment.

References

- Alfonso, M., Santiago, A., & Bassi, M. (2010). An Alternative Pathway into Teaching: Placing Top University Graduates in Vulnerable Schools in Chile. Inter-American Development Bank.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in education*, 5(1), 7-74.
- Brown, H. D., & Abeywickrama, P. (2004). *Language assessment. Principles and Classroom Practices*. White Plains, NY: Pearson Education.
- Centro de Medición de la Pontificia Universidad Católica de Chile (MIDE-UC). (2009a). Informe de Instrumentos para Evaluación de Aprendizajes, Habilidades Intelectuales e Intrapersonales – Programa Enseña Chile. Unpublished manuscript, Inter-American Development Bank.
- Centro de Medición de la Pontificia Universidad Católica de Chile (MIDE-UC). (2009b). Informe de Resultados de la Primera Aplicación de Pruebas SEPA – Enseña Chile. Unpublished manuscript, Inter-American Development Bank.
- Centro de Medición de la Pontificia Universidad Católica de Chile (MIDE-UC). (2009c). Sistema de Evaluación de Progreso del Aprendizaje (SEPA). <http://www.sepauc.cl/>.
- Center for Microdata, Economics Department, Universidad de Chile. (2011). Estudio sobre el Comportamiento Lector a Nivel Nacional. http://www.uchile.cl/documentos/estudio-sobre-el-comportamiento-lector-a-nivel-nacional_110593_2_2405.pdf.

- Cerrolaza, O., Cuadrado, C., Díaz, A. & Martín, M. (1999). El placer de aprender. Carabela 41, 23-33. SGEL, Madrid.
- Cimermanová, I. (2015). Using Comics with Novice EFL Readers to Develop Reading Literacy. *Procedia-Social and Behavioral Sciences*, 174, 2452-2459.
- Cope, B., & Kalantzis, M. (2009). A grammar of multimodality. *The International Journal of Learning*, 16(4), 361-426.
- Cook, T. D., Campbell, D. T., & Day, A. (1979). *Quasi-experimentation: Design & analysis issues for field settings* (Vol. 351). Boston: Houghton Mifflin.
- Evans, J. (2013). From comics, graphic novels and picturebooks to fusion texts: a new kid on the block!. *Education 3-13*, 41(2), 233-248.
- Fletcher, J. M. (2006). Measuring reading comprehension. *Scientific Studies of Reading*, 10(3), 323-330.
- Gee, J. P. (2004). *Situated language and learning: A critique of traditional schooling*. Psychology Press.
- Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge University Press.
- Kress, G. (2003). *Literacy in the new media age*. Psychology Press.
- León, J. A., & Pérez, O. (2003). Taxonomías y tipos de inferencias. JA León (Ed.), *Conocimiento y Discurso claves para inferir y comprender*, 45-66.
- McKenna, M. C., & Stahl, K. A. D. (2012). *Assessment for reading instruction*. Guilford Press.
- National Center for Education Statistics, U.S. Department of Education. (2007). *The Condition of Education 2007*.

- Nazari, B., Mansouri, S. (2014). Dynamic assessment versus static assessment: A study of reading comprehension ability in Iranian EFL learners. *Journal of Language and Linguistic Studies*, 10(2), 134-156.
- O'Reilly, T., Weeks, J., Sabatini, J., Halderman, L., Steinberg, J. (2014). Designing reading comprehension assessments for reading interventions: How a theoretically motivated assessment can serve as an outcome measure. *Educational Psychology Review*, 26(3), 403-424.
- Taut, S., Valencia, E., Palacios, D., Santelices, M. V., Jiménez, D., & Manzi, J. (2014). Teacher performance and student learning: linking evidence from two national assessment programmes. *Assessment in Education: Principles, Policy & Practice*, (ahead-of-print), 1-22.
- TERCE. (2015). Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación.
<http://www.unesco.org/new/fileadmin/MULTIMEDIA/FIELD/Santiago/pdf/TERCE-Cuadernillo2-Logros-aprendizaje-WEB.pdf>.
- Valerie, L., & Abed, F. (2013, January). Comic life + writing = motivated student writers: incorporating visual graphics to teach writing. In *Preparing Teachers to Teach Writing Using Technology* (pp. 163-178). ETC Press.

Table 1

Sample questions for each section

ID	Section	Difficulty Level	Question
1	1	Basic	In which city does the story begin?
7	2	Intermediate	Who is Passepartout's new boss?
14	3	Advanced	What is Mr. Fix's job?
16	4	Basic	Why does Fogg go to bed?
22	5	Intermediate	Where does the train stop?
30	6	Advanced	Why do they arrest Fogg and Passepartout?
31	7	Basic	Who discovers that the ship is being repaired?
37	8	Intermediate	Why does Passepartout fall asleep in the bar?
44	9	Advanced	Why did Fogg miss the boat to Yokohama?
46	10	Basic	Why does Passepartout join the circus?
53	11	Intermediate	Where do Fogg and Passepartout meet each other again?
60	12	Advanced	Why does Passepartout hit Fix?
61	13	Basic	How long was the train delayed by the bison?
67	14	Intermediate	Why does Fogg go to look for Passepartout?
74	15	Advanced	Whose idea is it to use wood from the boat for the stove?
76	16	Basic	Where does Fix arrest Fogg?
83	17	Intermediate	Why does Fogg think he lost the bet?
85	17	Advanced	Why did Fogg and Passepartout gain a day on their journey?

Note. This table includes one question per section, from a total of 85 questions.

Table 2

Analysis of the reliability index for each section

Section	Mean score if the section is removed	Variance of scores if the section is removed	Adjusted section-total correlation	Cronbach's alpha if the section is removed
Section 1	28.5025	84.756	0.206	0.830
Section 2	28.1357	81.199	0.234	0.831
Section 3	28.5377	76.815	0.408	0.822
Section 4	28.3116	76.458	0.521	0.815
Section 5	28.1206	77.400	0.486	0.817
Section 6	27.8794	78.238	0.502	0.817
Section 7	28.0302	77.151	0.504	0.816
Section 8	28.1558	77.021	0.452	0.819
Section 9	28.8090	78.731	0.375	0.823
Section 10	28.7739	80.600	0.343	0.825
Section 11	28.5075	77.170	0.425	0.821
Section 12	29.0352	78.458	0.385	0.823
Section 13	28.9749	76.025	0.522	0.815
Section 14	28.6382	74.384	0.570	0.812
Section 15	28.6181	74.419	0.559	0.812
Section 16	28.6181	74.712	0.526	0.814
Section 17	29.0955	80.218	0.251	0.832

Note. This analysis was made to see whether it was necessary to remove any of the sections.

Table 3

Results of the linear regressions based on the scores from the instrument

Scores on the Instrument with	B	P value	R²
Language Arts Grade	0.3	< 0.001	0.241
Percentage score on SEPA test	0.105	< 0.001	0.315

Table 4

Results from the two-way ANOVA test, comparing the results obtained by girls and boys

Category	F	gl	P value
Language Arts Grade	11.13	1/180	0.00103
Score on Instrument	4.757	1/180	0.0305
Time (min) on Instrument	3.56	1/180	0.0608
Percentage Score on SEPA Test	6.542	1/180	0.0114

Table 5

Results for the scores obtained on the instrument, as well as time spent completing the test

	Score on the Instrument (1.0 to 7.0, 7.0 highest score)	Time spent on the instrument (min)
Minimum	2.20	13.22
Maximum	7.00	45.43
Average	4.65	25.02

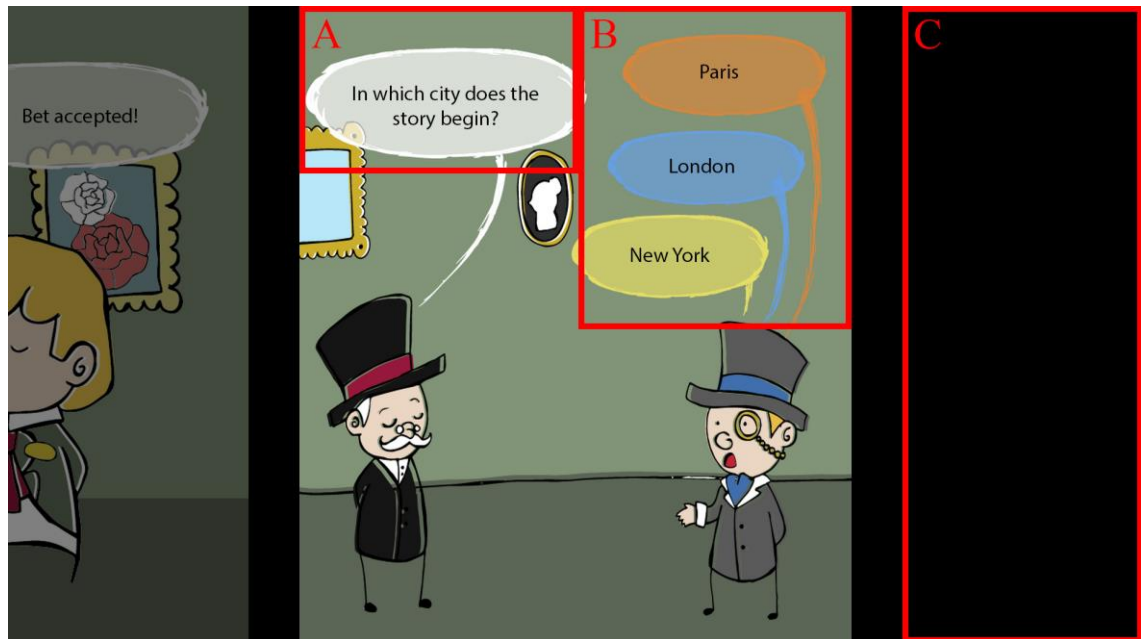
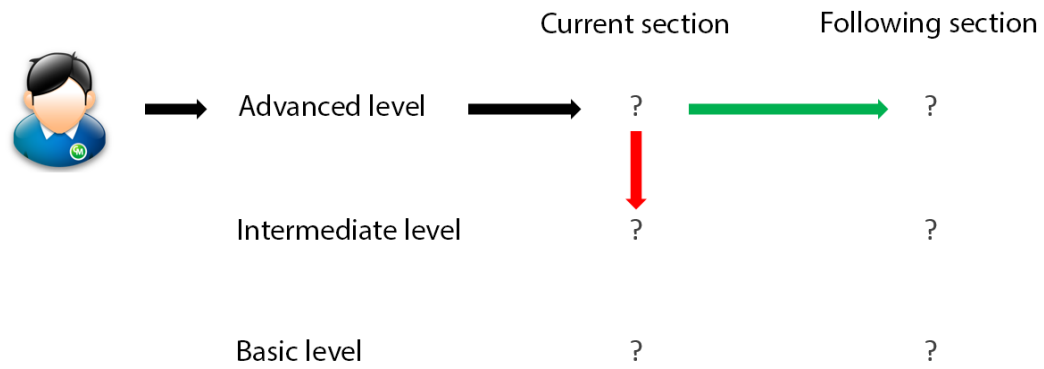
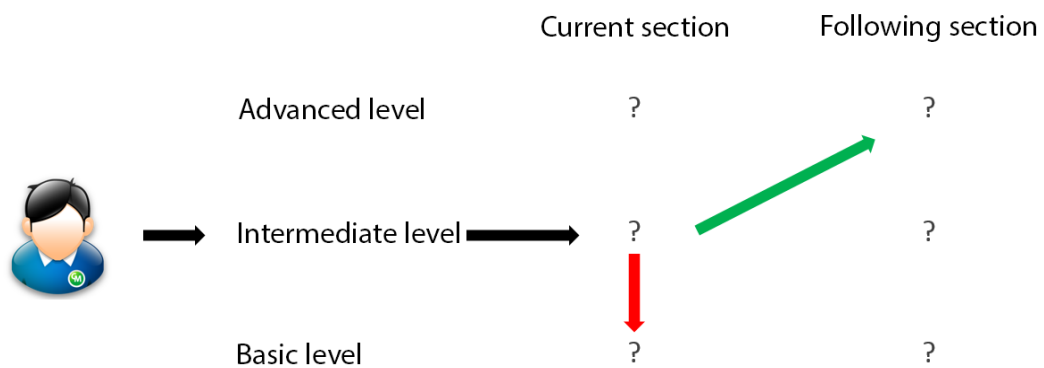


Figure 1. Question at the end of a section. The “next panel” area appears blank.

Error = 0



Error = 0



Error = 0

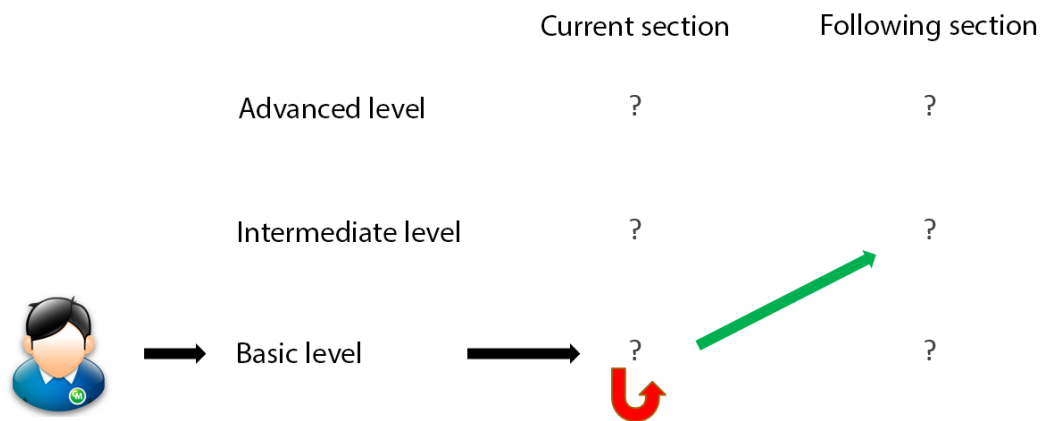
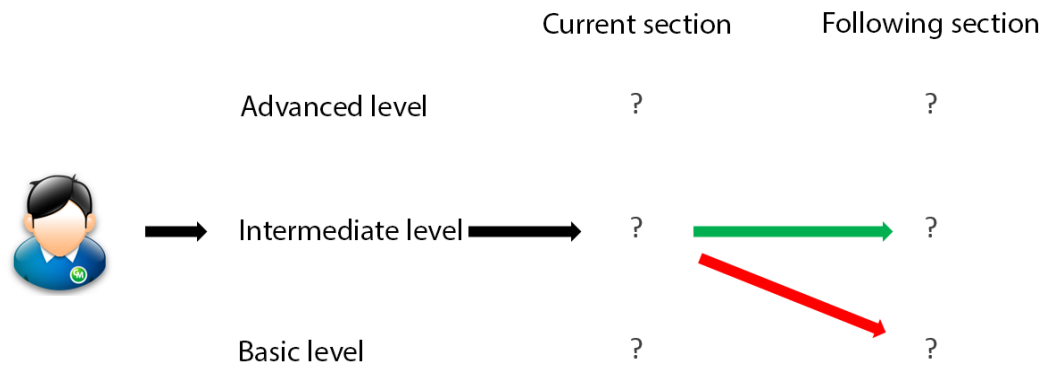


Figure 2. Detailed flow between difficulty levels when there are no accumulated errors.

Error = 1



Error = 1

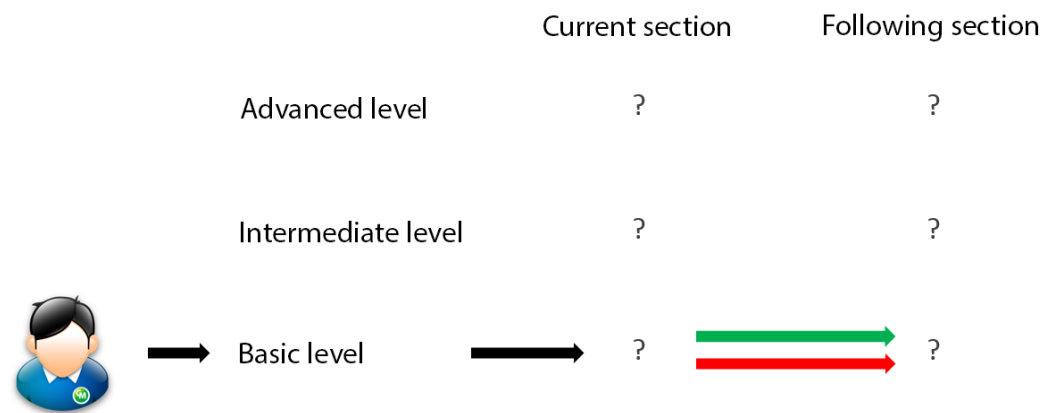


Figure 3. Detailed flow between difficulty levels, where accumulated errors is equal to one.



Figure 4. Image of the screen when using the instrument.

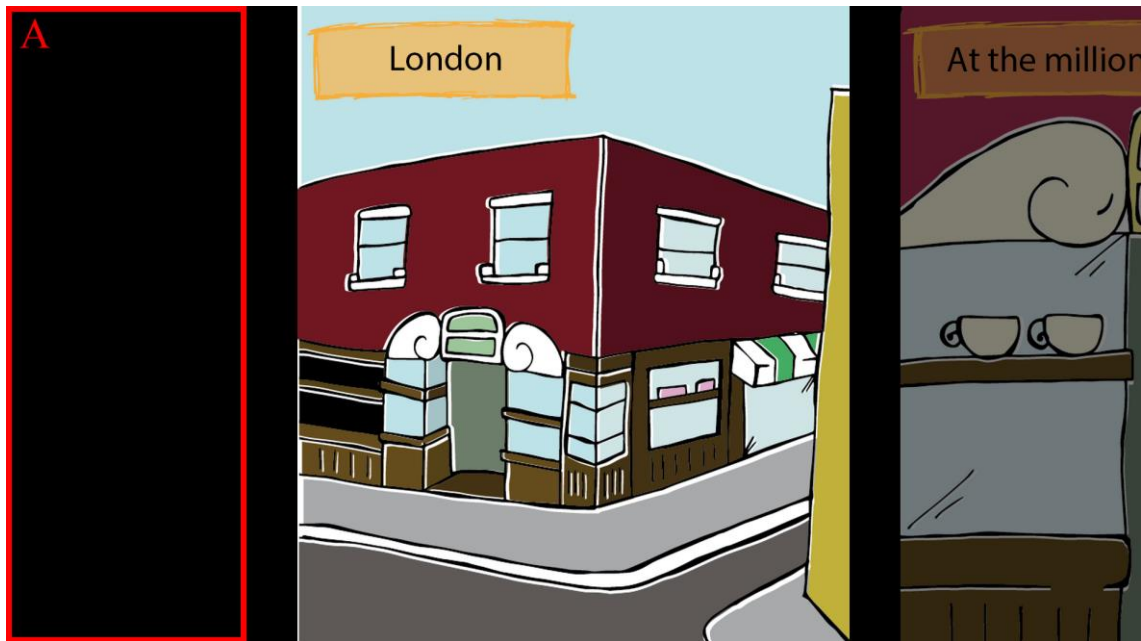


Figure 5. The “previous panel” area is empty at the beginning of a section.

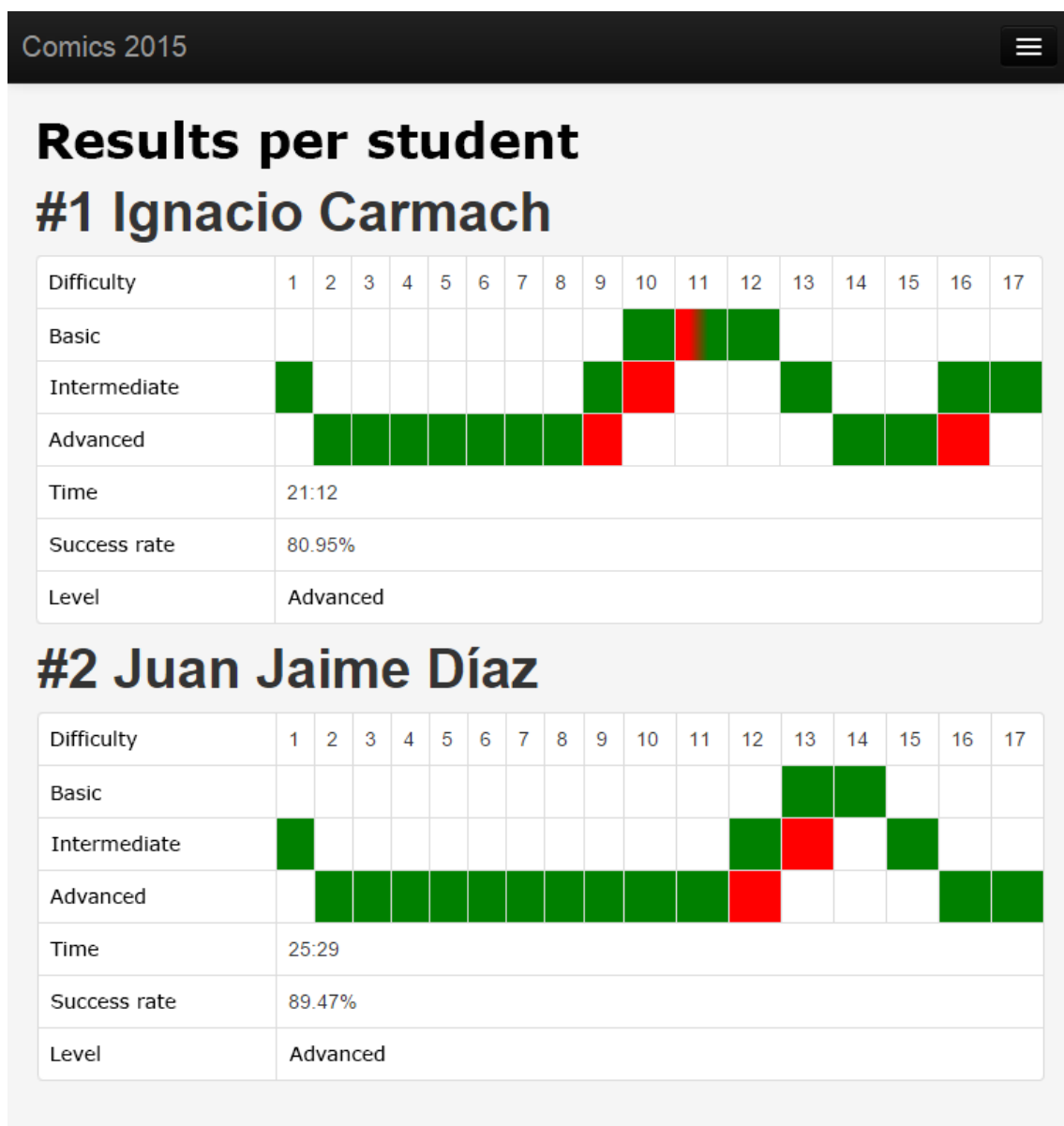


Figure 6. Example of progress report from the assessment.

ANEXO C: PREGUNTAS DISEÑADAS PARA EL INSTRUMENTO

ID	Sección	Nivel de dificultad	Opción	Pregunta
1	1	Básico	1	¿En qué ciudad comienza la historia?
2	1	Intermedio	1	¿En qué ciudad comienza la historia?
3	1	Intermedio	2	¿Por qué es difícil atrapar al ladrón?
4	1	Avanzado	1	¿Por qué es difícil atrapar al ladrón?
5	1	Avanzado	2	¿Qué podría retrasar el viaje alrededor del mundo?
6	2	Básico	1	¿Quién es el nuevo jefe de Picaporte?
7	2	Intermedio	1	¿Quién es el nuevo jefe de Picaporte?
8	2	Intermedio	2	¿Cuál es la primera tarea de Picaporte?
9	2	Avanzado	1	¿Cuál es la primera tarea de Picaporte?
10	2	Avanzado	2	¿Quién es sospechoso de ser el ladrón del banco?
11	3	Básico	1	¿Cuál es la primera parada del viaje?
12	3	Intermedio	1	¿Cuál es la primera parada del viaje?
13	3	Intermedio	2	¿Qué profesión tiene el señor Fix?
14	3	Avanzado	1	¿Qué profesión tiene el señor Fix?
15	3	Avanzado	2	¿Por qué sospechan que el señor Fogg es el ladrón?
16	4	Básico	1	¿Por qué Fogg se va a dormir?
17	4	Intermedio	1	¿Por qué Fogg se va a dormir?
18	4	Intermedio	2	¿En qué medio de transporte viajaron a Bombay?
19	4	Avanzado	1	¿En qué medio de transporte viajaron a Bombay?

20	4	Avanzado	2	¿Por qué al detective le parece sospechoso Picaporte?
21	5	Básico	1	¿Dónde se detiene el tren?
22	5	Intermedio	1	¿Dónde se detiene el tren?
23	5	Intermedio	2	¿Cómo deciden viajar para llegar más rápido a Calcuta?
24	5	Avanzado	1	¿Cómo deciden viajar para llegar más rápido a Calcuta?
25	5	Avanzado	2	¿Por qué Fogg y Picaporte decidieron continuar el viaje en elefante?
26	6	Básico	1	¿Quiénes detienen a Fogg y Picaporte?
27	6	Intermedio	1	¿Quiénes detienen a Fogg y Picaporte?
28	6	Intermedio	2	¿Qué hizo Fogg para quedar libre del arresto?
29	6	Avanzado	1	¿Qué hizo Fogg para quedar libre del arresto?
30	6	Avanzado	2	¿Por qué razón detienen a Fogg y a Picaporte?
31	7	Básico	1	¿Quién descubre que el barco está en reparaciones?
32	7	Intermedio	1	¿Quién descubre que el barco está en reparaciones?
33	7	Intermedio	2	¿Cuál es el siguiente destino del viaje de Fogg y Picaporte?
34	7	Avanzado	1	¿Cuál es el siguiente destino del viaje de Fogg y Picaporte?
35	7	Avanzado	2	¿En qué ciudad fue la tormenta?
36	8	Básico	1	¿Por qué Picaporte se queda dormido en el bar?
37	8	Intermedio	1	¿Por qué Picaporte se queda dormido en el bar?
38	8	Intermedio	2	¿Por qué Picaporte tiene tiempo para tomar algo en

				el bar?
39	8	Avanzado	1	¿Por qué Picaporte tiene tiempo para tomar algo en el bar?
40	8	Avanzado	2	¿Por qué Fix envenena a Picaporte?
41	9	Básico	1	¿Cómo piensa Fogg llegar a Yokohama?
42	9	Intermedio	1	¿Cómo piensa Fogg llegar a Yokohama?
43	9	Intermedio	2	¿Por qué Fogg perdió el barco a Yokohama?
44	9	Avanzado	1	¿Por qué Fogg perdió el barco a Yokohama?
45	9	Avanzado	2	¿Por qué Fogg aborda un barco a Nagasaki y Shangai?
46	10	Básico	1	¿Para qué se une Picaporte al circo?
47	10	Intermedio	1	¿Para qué se une Picaporte al circo?
48	10	Intermedio	2	¿En qué barco despierta Picaporte?
49	10	Avanzado	1	¿En qué barco despierta Picaporte?
50	10	Avanzado	2	¿Qué pista tuvo el empleado del barco para saber que Picaporte era un pasajero a Yokohama?
51	11	Básico	1	¿Por qué Picaporte no le cuenta a Fogg sobre el detective que lo está siguiendo?
52	11	Intermedio	1	¿Por qué Picaporte no le cuenta a Fogg sobre el detective que lo está siguiendo?
53	11	Intermedio	2	¿Dónde se reencuentran Fogg y Picaporte?
54	11	Avanzado	1	¿Dónde se reencuentran Fogg y Picaporte?
55	11	Avanzado	2	Fogg y Picaporte están en la mitad del camino. Según el plan inicial, ¿cuántos días deberían faltar para llegar a su destino?

56	12	Básico	1	¿Dónde planea Fix capturar a Fogg?
57	12	Intermedio	1	¿Dónde planea Fix capturar a Fogg?
58	12	Intermedio	2	¿Dónde tomarán Fogg y Picaporte el próximo tren?
59	12	Avanzado	1	¿Dónde tomarán Fogg y Picaporte el próximo tren?
60	12	Avanzado	2	¿Por qué Picaporte golpea a Fix?
61	13	Básico	1	¿Cuánto se retrasó el tren por causa de los bisontes?
62	13	Intermedio	1	¿Cuánto se retrasó el tren por causa de los bisontes?
63	13	Intermedio	2	¿Por qué está enojado Fogg?
64	13	Avanzado	1	¿Cuánto se retrasó el tren por causa de los bisontes?
65	13	Avanzado	2	¿Por qué está enojado Fogg?
66	14	Básico	1	¿Por qué Fogg sale a buscar a Picaporte?
67	14	Intermedio	1	¿Por qué Fogg sale a buscar a Picaporte?
68	14	Intermedio	2	¿Por qué piensa Fogg que Picaporte está herido?
69	14	Avanzado	1	¿Por qué piensa Fogg que Picaporte está herido?
70	14	Avanzado	2	¿Por qué Fogg y Picaporte deciden tomar un tren de carga?
71	15	Básico	1	¿Cuánto dinero ofrece Fogg al capitán del barco para que cambie de destino?
72	15	Intermedio	1	¿Cuánto dinero ofrece Fogg al capitán del barco para que cambie de destino?
73	15	Intermedio	2	¿A quién se le ocurre usar madera del mismo barco para la caldera?

74	15	Avanzado	1	¿A quién se le ocurre usar madera del mismo barco para la caldera?
75	15	Avanzado	2	Fogg debe pagar dos veces al capitán del barco para llegar a su destino, ¿Qué pagan cada vez?
76	16	Básico	1	¿Dónde arresta Fix a Fogg?
77	16	Intermedio	1	¿Dónde arresta Fix a Fogg?
78	16	Intermedio	2	¿Qué hace Fix para reparar el daño causado a Fogg?
79	16	Avanzado	1	¿Qué hace Fix para reparar el daño causado a Fogg?
80	16	Avanzado	2	¿Por qué Fix se disculpa con Fogg?
81	17	Básico	1	¿Cuándo salieron Fogg y Picaporte de Londres?
82	17	Intermedio	1	¿Cuándo salieron Fogg y Picaporte de Londres?
83	17	Intermedio	2	¿Por qué Fogg cree que perdió la apuesta?
84	17	Avanzado	1	¿Por qué Fogg cree que perdió la apuesta?
85	17	Avanzado	2	¿Por qué Fogg y Picaporte ganaron un día en su viaje?

**ANEXO D: RESULTADOS DE LOS DISTINTOS INSTRUMENTOS
UTILIZADOS EN LA INVESTIGACIÓN**

Id	Género	Colegio	Tipo de Colegio	Resultados Test Dinámico	Notas Lenguaje y Comunicación	Tiempo Test Dinámico (seg)	Sepa Exp	Sepa Imp	Sepa Sgl	Sepa Sce	Sepa CrI	Sepa Global
1	0	1	0	5.80	5.80	1690	0.33	0.50	0.29	0.29	0.75	0.41
2	0	1	0	5.92	6.60	1536	1.00	0.67	0.86	1.00	0.75	0.85
3	0	1	0	5.08	5.60	1965	0.56	0.67	0.14	0.86	0.75	0.59
4	1	1	0	4.00	5.00	2259	0.11	0.33	0.14	0.29	0.25	0.23
5	1	1	0	4.72	5.60	1431	0.67	0.50	0.29	0.29	0.25	0.44
6	1	1	0	4.48	6.00	1884	0.33	0.42	0.29	0.86	0.50	0.46
7	1	1	0	5.20	5.70	1520	0.33	0.42	0.57	0.57	0.50	0.46
8	1	1	0	6.16	5.50	1440	0.89	0.83	0.71	0.86	0.75	0.82
9	1	1	0	4.60	5.30	1850	0.67	0.58	0.29	0.86	0.75	0.62
10	1	1	0	2.80	6.10	1770	0.67	0.75	0.57	0.71	0.25	0.64
11	1	1	0	7.00	5.30	1351	0.89	0.75	0.71	1.00	0.75	0.82
12	0	1	0	5.68	6.00	2336	0.67	0.50	0.29	0.71	1.00	0.59
13	0	1	0	4.00	6.20	1064	0.89	0.83	0.71	1.00	1.00	0.87
14	0	1	0	5.08	5.70	1461	0.44	0.42	0.43	0.14	0.50	0.38
15	1	1	0	4.96	5.50	1651	0.56	0.17	0.43	0.71	0.75	0.46
16	1	1	0	5.20	5.80	1163	0.89	0.75	0.43	1.00	0.75	0.77
17	0	1	0	4.84	5.50	1991	0.44	0.25	0.43	0.43	0.25	0.36
18	0	1	0	6.76	5.90	1995	0.78	0.83	0.86	0.71	0.75	0.79
19	0	1	0	5.44	6.30	2033	0.67	0.58	0.71	0.57	0.50	0.62
20	0	2	0	4.12	5.90	1265	0.13	0.50	0.43	0.67	0.25	0.40
21	0	2	0	5.68	6.50	1429	1.00	0.70	0.57	1.00	0.75	0.80

22	0	2	0	6.64	5.30	2038	1.00	0.40	0.57	1.00	0.00	0.63
23	0	2	0	3.40	4.50	1695	0.38	0.40	0.57	0.17	0.00	0.34
24	0	2	0	4.36	4.60	1975	0.25	0.20	0.00	0.67	0.25	0.26
25	0	2	0	4.48	5.50	2238	0.63	0.40	0.57	1.00	0.25	0.57
26	0	2	0	5.68	4.90	1243	0.63	0.30	0.43	0.67	0.25	0.46
27	0	2	0	4.12	6.60	1428	0.75	0.70	0.57	1.00	0.25	0.69
28	0	2	0	3.76	5.20	1478	0.75	0.40	0.57	0.83	0.75	0.63
29	0	2	0	4.72	5.20	1623	0.88	0.40	0.71	1.00	0.50	0.69
30	0	2	0	4.12	5.90	1024	0.63	0.60	0.71	0.83	0.50	0.66
31	0	2	0	2.92	5.00	1988	0.38	0.60	0.43	0.67	0.50	0.51
32	0	2	0	3.04	5.70	1299	0.50	0.60	0.00	0.67	0.25	0.43
33	0	2	0	4.60	5.00	1925	0.13	0.30	0.43	0.33	0.00	0.26
34	0	2	0	3.88	5.40	2188	0.63	0.40	0.43	0.83	0.50	0.54
35	0	2	0	3.16	5.00	1287	0.38	0.40	0.14	0.33	0.25	0.31
36	0	2	0	3.76	4.00	1643	0.63	0.10	0.43	0.50	0.00	0.34
37	0	2	0	2.68	4.80	1257	0.38	0.50	0.14	0.67	0.25	0.40
38	0	2	0	4.72	4.50	1358	0.63	0.50	0.29	1.00	0.75	0.60
39	0	2	0	5.44	5.50	1370	0.50	0.60	0.43	1.00	0.50	0.60
40	0	2	0	2.20	4.90	1385	0.63	0.40	0.43	0.83	0.50	0.54
41	1	2	0	5.68	5.20	1365	0.50	0.00	0.43	0.50	0.50	0.34
42	1	2	0	3.64	4.20	1397	0.25	0.60	0.00	0.33	0.50	0.34
43	1	2	0	2.92	4.30	2198	0.25	0.30	0.14	0.00	0.25	0.20
44	1	2	0	2.56	4.00	477	0.38	0.30	0.00	0.50	0.75	0.34
45	1	2	0	5.80	5.30	2087	0.75	0.40	0.71	0.67	0.50	0.60
46	1	2	0	3.76	4.90	1188	0.63	0.60	0.71	0.83	0.50	0.66
47	1	2	0	2.32	4.70	2045	0.25	0.40	0.29	0.17	0.25	0.29
48	1	2	0	3.52	5.20	1414	0.13	0.60	0.29	0.67	0.00	0.37

49	1	2	0	2.44	4.40	1668	0.38	0.80	0.29	0.17	0.25	0.43
50	1	2	0	3.88	6.10	1124	0.88	0.30	0.71	0.67	0.25	0.57
51	1	2	0	3.16	5.20	1134	0.38	0.30	0.29	0.33	0.00	0.29
52	1	2	0	4.12	4.80	1508	0.38	0.60	0.14	0.33	0.50	0.40
53	1	2	0	3.04	5.50	2030	0.25	0.30	0.29	0.33	0.00	0.26
54	1	2	0	5.92	5.80	1413	0.63	0.40	0.71	1.00	0.50	0.63
55	1	2	0	2.92	5.40	1764	0.38	0.50	0.43	0.67	0.50	0.49
56	1	2	0	6.40	6.60	1440	0.88	0.70	1.00	1.00	1.00	0.89
57	1	2	0	2.92	5.90	1248	0.38	0.30	0.43	0.50	0.50	0.40
58	1	2	0	3.76	5.90	1487	0.25	0.50	0.29	0.33	0.25	0.34
59	1	2	0	4.96	5.80	1106	0.25	0.20	0.14	0.33	0.25	0.23
60	1	2	0	2.56	4.90	1471	0.13	0.20	0.14	0.33	0.50	0.23
61	1	2	0	4.72	4.80	1359	0.75	0.90	0.71	1.00	0.75	0.83
62	1	2	0	4.12	6.00	1224	0.25	0.40	0.14	0.50	0.25	0.31
63	1	2	0	2.92	4.80	1985	1.00	0.50	0.14	0.83	0.25	0.57
64	1	2	0	6.40	6.30	2034	1.00	0.60	1.00	1.00	0.75	0.86
65	1	2	0	5.56	6.20	1435	0.88	0.80	0.86	1.00	0.75	0.86
66	1	2	0	5.68	5.50	1955	0.75	0.30	0.57	0.67	0.50	0.54
67	1	2	0	4.36	5.90	2172	0.50	0.30	0.29	0.17	0.25	0.31
68	1	2	0	4.84	5.10	1685	0.63	0.60	0.29	0.33	0.75	0.51
69	1	2	0	4.48	4.50	2068	0.63	0.30	0.71	1.00	0.75	0.63
70	1	2	0	4.36	4.70	1527	0.25	0.40	0.29	0.33	0.25	0.31
71	1	2	0	3.04	5.20	2034	0.13	0.40	0.14	0.33	0.00	0.23
72	1	2	0	4.24	4.40	1845	0.50	0.60	0.43	0.67	0.00	0.49
73	1	2	0	4.72	5.20	1026	0.50	0.50	0.29	0.50	0.50	0.46
74	1	2	0	3.76	4.90	1402	0.38	0.50	0.57	0.17	0.75	0.46
75	1	2	0	3.40	4.00	1381	0.25	0.30	0.29	0.67	0.25	0.34

76	1	2	0	4.36	4.60	930	0.50	0.50	0.57	0.67	0.00	0.49
77	1	2	0	3.88	4.90	1369	0.50	0.40	0.57	0.67	0.50	0.51
78	1	2	0	4.00	6.50	1322	1.00	0.70	0.14	1.00	0.75	0.71
79	0	3	1	5.32	6.20	2074	0.50	0.30	0.71	1.00	0.50	0.57
80	0	3	1	5.92	6.20	2019	0.75	0.60	0.57	1.00	0.75	0.71
81	0	3	1	5.20	6.50	1899	0.88	0.60	0.71	0.83	0.75	0.74
82	0	3	1	4.72	5.90	1457	0.38	0.40	0.57	0.67	0.50	0.49
83	0	3	1	5.44	6.60	1445	0.88	0.50	0.43	1.00	0.75	0.69
84	0	3	1	2.92	6.50	1710	0.75	0.70	0.57	1.00	1.00	0.77
85	0	3	1	6.04	6.30	1237	0.88	0.80	0.86	1.00	0.50	0.83
86	0	3	1	5.92	7.00	1908	0.75	0.80	0.86	1.00	1.00	0.86
87	0	3	1	6.52	6.80	1561	1.00	1.00	0.86	1.00	1.00	0.97
88	0	3	1	5.56	6.70	1562	1.00	0.70	0.86	1.00	1.00	0.89
89	0	3	1	3.52	6.00	1473	0.63	0.40	0.43	0.83	0.25	0.51
90	0	3	1	4.72	6.60	1433	0.88	0.80	0.57	1.00	0.50	0.77
91	0	3	1	5.08	6.60	1717	0.50	0.40	0.57	0.83	0.50	0.54
92	0	3	1	4.60	6.70	1158	1.00	0.70	1.00	1.00	1.00	0.91
93	0	3	1	5.92	6.60	1725	0.88	0.80	0.86	1.00	0.75	0.86
94	0	3	1	4.48	5.80	1115	0.88	0.90	0.71	1.00	0.25	0.80
95	0	3	1	5.20	5.90	2035	0.50	0.70	0.71	0.67	0.25	0.60
96	0	3	1	4.96	6.60	1421	0.88	0.60	0.71	1.00	1.00	0.80
97	0	3	1	5.08	6.50	1502	0.75	0.60	0.57	1.00	0.50	0.69
98	0	3	1	5.20	7.00	1080	0.63	0.70	1.00	0.83	0.75	0.77
99	1	3	1	5.44	5.90	1376	1.00	0.90	1.00	1.00	1.00	0.97
100	1	3	1	5.92	6.00	1153	0.75	0.70	1.00	1.00	0.75	0.83
101	1	3	1	4.00	6.10	1070	0.88	0.80	0.71	1.00	0.75	0.83
102	1	3	1	2.68	5.80	850	0.88	0.80	0.43	0.83	0.25	0.69

103	1	3	1	5.32	6.30	1338	0.75	0.70	1.00	0.83	1.00	0.83
104	1	3	1	5.08	5.90	1243	0.75	0.80	0.71	0.83	0.50	0.74
105	1	3	1	4.36	5.90	1544	0.63	0.30	0.71	1.00	0.25	0.57
106	1	3	1	4.72	6.60	1099	0.88	0.50	0.86	1.00	0.75	0.77
107	1	3	1	4.48	6.20	1553	0.88	0.70	0.71	1.00	1.00	0.83
108	1	3	1	5.08	5.80	1365	0.63	0.20	0.43	1.00	0.75	0.54
109	1	3	1	5.32	6.20	1033	0.88	0.60	1.00	1.00	0.75	0.83
110	1	3	1	4.84	5.90	1329	0.63	0.70	0.43	0.83	1.00	0.69
111	1	3	1	4.84	6.30	1192	0.88	0.90	0.71	0.83	0.75	0.83
112	1	3	1	4.24	6.10	1407	0.63	0.60	0.43	0.33	0.50	0.51
113	1	3	1	2.20	5.00	1000	0.38	0.30	0.14	0.67	0.50	0.37
114	1	3	1	5.68	6.80	1014	1.00	0.90	1.00	1.00	0.75	0.94
115	1	3	1	3.88	4.80	1166	0.13	0.20	0.14	0.67	1.00	0.34
116	1	3	1	6.28	6.20	1529	1.00	0.60	0.71	1.00	1.00	0.83
117	1	3	1	4.36	6.00	1526	0.50	0.50	0.43	0.83	0.25	0.51
118	1	3	1	5.56	6.20	1207	0.75	0.40	0.43	0.83	0.00	0.51
119	1	3	1	4.00	5.80	1307	0.88	0.80	0.71	1.00	1.00	0.86
120	1	3	1	3.04	5.10	1285	0.38	0.20	0.29	0.00	0.25	0.23
121	1	3	1	4.72	6.20	893	0.63	0.50	0.71	1.00	0.25	0.63
122	1	3	1	4.72	6.70	1050	0.88	0.80	1.00	1.00	0.75	0.89
123	1	3	1	5.56	5.90	1786	1.00	0.70	0.86	0.83	0.25	0.77
124	1	3	1	3.52	5.30	1054	0.25	0.00	0.00	0.17	0.50	0.14
125	1	3	1	5.44	6.80	1075	0.88	0.80	0.86	0.83	1.00	0.86
126	0	3	1	5.44	6.70	1781	0.88	0.40	1.00	1.00	0.50	0.74
127	0	3	1	5.68	6.60	1286	0.63	0.60	0.57	0.67	0.50	0.60
128	0	3	1	6.28	6.70	1446	0.75	0.80	0.86	1.00	1.00	0.86
129	0	3	1	6.04	6.20	1560	0.75	0.90	0.71	1.00	1.00	0.86

130	0	3	1	4.84	5.80	2029	0.88	0.30	0.57	0.67	0.75	0.60
131	0	3	1	4.72	6.10	1653	0.88	0.50	0.57	0.83	0.75	0.69
132	0	3	1	4.24	6.70	1278	1.00	0.60	1.00	1.00	0.75	0.86
133	0	3	1	5.56	6.50	1502	1.00	0.80	0.71	1.00	0.75	0.86
134	0	3	1	6.52	6.50	1398	0.88	0.50	0.71	1.00	0.50	0.71
135	0	3	1	3.40	5.40	1287	0.50	0.40	0.71	1.00	0.75	0.63
136	0	3	1	3.52	6.60	1703	1.00	0.70	0.57	1.00	0.75	0.80
137	0	3	1	3.76	5.90	1615	0.63	0.70	0.43	0.50	0.25	0.54
138	0	3	1	4.24	6.20	1307	0.25	0.50	0.00	0.67	0.25	0.34
139	0	3	1	5.56	6.20	1196	1.00	0.60	0.86	1.00	0.75	0.83
140	0	3	1	5.08	6.20	1437	0.75	0.70	0.29	0.50	0.25	0.54
141	0	3	1	5.20	6.00	1390	0.63	0.40	0.71	1.00	0.00	0.57
142	0	3	1	4.96	6.40	1516	0.75	0.70	0.86	1.00	0.25	0.74
143	0	3	1	4.96	5.60	2405	0.50	0.20	0.43	0.17	0.50	0.34
144	0	3	1	4.48	6.20	1379	0.75	0.40	0.86	0.67	0.75	0.66
145	0	3	1	4.72	6.20	1143	0.88	0.60	1.00	1.00	0.50	0.80
146	0	3	1	5.80	6.60	1552	0.75	0.80	1.00	1.00	0.75	0.86
147	0	3	1	5.68	5.10	1630	0.38	0.50	0.29	0.83	0.25	0.46
148	0	3	1	6.28	6.60	1147	0.88	0.90	0.71	1.00	0.75	0.86
149	0	3	1	5.20	6.70	1075	0.88	0.90	0.71	1.00	1.00	0.89
150	0	3	1	3.40	5.80	1223	0.75	0.60	0.57	1.00	0.50	0.69
151	0	3	1	3.64	6.50	1329	1.00	0.50	0.43	1.00	0.75	0.71
152	0	3	1	4.36	6.00	1395	0.75	0.50	0.57	0.83	0.25	0.60
153	0	3	1	4.48	6.50	1868	1.00	0.70	0.86	0.83	0.50	0.80
154	0	3	1	4.84	6.70	1028	1.00	0.80	1.00	1.00	0.50	0.89
155	0	3	1	5.20	5.50	1313	0.75	0.90	0.86	1.00	1.00	0.89
156	0	3	1	6.28	6.40	1011	0.75	0.70	0.71	0.83	0.75	0.74

157	1	3	1	4.72	5.90	1584	0.75	0.50	0.29	0.67	0.50	0.54
158	1	3	1	4.84	6.30	2175	0.38	0.20	0.29	0.00	0.25	0.23
159	1	3	1	6.52	5.80	1159	0.88	0.80	1.00	1.00	0.75	0.89
160	1	3	1	6.76	6.60	1333	1.00	0.90	1.00	1.00	0.75	0.94
161	1	3	1	5.92	6.40	1274	1.00	0.70	0.71	1.00	0.50	0.80
162	1	3	1	5.44	6.20	1573	0.75	0.30	0.57	0.67	0.75	0.57
163	1	3	1	3.76	5.40	2726	0.25	0.20	0.29	0.17	0.25	0.23
164	1	3	1	4.60	6.50	996	0.63	0.70	0.57	0.83	0.50	0.66
165	1	3	1	4.60	5.60	1446	0.88	0.70	0.71	0.83	0.50	0.74
166	1	3	1	5.56	6.70	1024	1.00	0.60	0.86	0.83	1.00	0.83
167	1	3	1	4.96	6.30	793	1.00	0.70	1.00	0.83	0.50	0.83
168	1	3	1	5.20	5.50	1451	1.00	0.60	0.71	1.00	0.50	0.77
169	1	3	1	2.92	5.80	962	0.38	0.50	0.71	1.00	0.00	0.54
170	1	3	1	5.32	6.70	947	0.63	0.80	0.86	0.83	0.75	0.77
171	1	3	1	2.92	5.60	1435	0.38	0.50	0.57	0.67	0.00	0.46
172	1	3	1	5.20	6.00	1378	0.75	0.20	0.71	0.83	0.75	0.60
173	1	3	1	3.52	5.70	1823	0.38	0.60	0.86	1.00	0.25	0.63
174	1	3	1	3.88	5.80	1720	0.75	0.60	0.43	1.00	0.25	0.63
175	1	3	1	6.40	6.20	1519	0.63	0.60	1.00	1.00	0.75	0.77
176	1	3	1	5.32	6.60	1263	0.88	0.50	0.43	0.83	0.75	0.66
177	1	3	1	4.36	5.80	1423	0.38	0.50	0.86	0.83	0.50	0.60
178	1	3	1	3.16	6.30	1637	0.75	0.40	0.71	0.83	0.25	0.60
179	1	3	1	4.24	6.40	1763	0.50	0.50	0.14	0.50	0.50	0.43
180	1	3	1	6.16	6.00	1285	0.75	0.80	1.00	0.67	0.75	0.80
181	1	3	1	3.88	5.10	2459	0.50	0.20	0.29	0.33	0.75	0.37
182	1	3	1	4.96	5.60	1018	0.25	0.50	0.57	0.67	0.75	0.51