

iSTART-E: Desarrollando un tutor inteligente para la comprensión lectora de estudiantes de  
habla hispana

Christian M. Soto, Danielle S. McNamara, Matthew E. Jacovina, Erica L. Snow, Jianmin Dai,  
Laura K. Allen, Cecile A. Perret, Amy M. Johnson, Devin G. Russell.

*Cualquier consulta sobre este artículo dirigirla a Christian Soto:*  
*christiansotofajardo@gmail.com*  
Institute for the Science of Teaching and Learning. Arizona State University

### *Descripción del programa iSTART*

El programa iSTART en español significa “Entrenamiento interactivo en estrategias para pensar y leer activamente”. El iSTART-2 es la última versión de este programa, que es un sistema tutorial inteligente (STI) basado en juegos que está diseñado para mejorar las habilidades de comprensión lectora (Jackson & McNamara, 2013; McNamara, Levinstein & Boonthum, 2004; Snow, Allen, Jacobina & McNamara, 2015). En concreto, el iSTART-2 se focaliza en la enseñanza de estrategias de auto-explicación que ayudan a mejorar la comprensión de textos de ciencia. La investigación sugiere que muchos lectores comprenden mejor los textos cuando pueden darse explicaciones a sí mismos a medida que los leen (Chi, de Leeuw, Chiu & Lavancher, 1994). Por ejemplo, al leer un texto sobre la evaporación, un lector puede generar una auto-explicación que relacione el texto con sus propias experiencias personales (por ejemplo, "*esto es lo que sucede cuando hago hervir los espaguetis y sale el vapor de la olla*"). De este modo *las auto-explicaciones* pueden permitir a los estudiantes construir modelos mentales más complejos que integren la información leída con los conocimientos previos. Sin embargo, no todos los estudiantes tienen el mismo éxito en explicarse a sí mismos los textos y pueden tener dificultades para generar explicaciones significativas que promuevan el entendimiento profundo de lo que leen. Por esta razón fue desarrollado el entrenamiento en estrategias de auto-explicación para ayudar a solucionar este problema (McNamara, 2004). El programa iSTART-2 y sus predecesores (iSTART y el iSTART-ME -mejorado motivacionalmente-) fueron diseñados para ofrecer entrenamiento en auto-explicación a través de un programa informático que proporciona retroalimentación adaptativa automatizada. El iSTART-2 presenta una fase de lección y una fase de práctica. Durante la fase de lección los estudiantes ven una serie de videos en los que un agente pedagógico proporciona instrucción en relación a cinco estrategias de auto-explicación. Estas

estrategias instruyen a los estudiantes para monitorear su comprensión, parafrasear el contenido del texto, predecir lo que más tarde abordará el texto, relacionar los conocimientos previos con el contenido del texto y realizar puentes de conexión entre diferentes partes del texto. Estos cinco videos de lección proporcionan a los estudiantes definiciones y ejemplos de las estrategias seguido de preguntas que ponen a prueba la comprensión de las mismas. Después de completar estas lecciones, los estudiantes ven un video que resume las estrategias y finalmente un video de demostración que ofrece ejemplos de cómo combinar las cinco estrategias de auto-explicación al leer un texto en particular.

Después de completar la fase de lección los estudiantes se dirigen a la interfaz de práctica del iSTART-2 en la que tienen acceso a varias características del sistema. En los juegos de identificación a los estudiantes se les presenta un texto, una oración meta y una auto-explicación. La tarea de los estudiantes consiste en seleccionar cuál de las cinco estrategias de auto-explicación se utiliza para generar la auto-explicación que se presenta. En los juegos generativos, a los estudiantes se les presenta un texto, una oración meta y se les pide escribir su propia explicación. Las auto-explicaciones de los estudiantes se califican automáticamente utilizando técnicas de procesamiento de lenguaje natural (PLN); puntuaciones automatizadas que son comparables con calificaciones humanas (Jackson, Guess, & McNamara, 2010). En general, las puntuaciones más bajas se dan a auto-explicaciones que son demasiado breves o irrelevantes, mientras que las puntuaciones más altas se dan a las auto-explicaciones que incorporan palabras claves del texto e información relevante externa al texto leído relacionada con el contenido del mismo (lo que sugiere que los estudiantes están incorporando conocimientos “*más allá del texto*”). En los juegos generativos los estudiantes reciben feedback acerca de sus auto-explicaciones, y específicamente en el juego “*coached practice*” el estudiante recibe información adicional para mejorar sus auto-explicaciones y revisar su desempeño. El algoritmo asociado a las

técnicas de PLN determina la retroalimentación que reciben los estudiantes. En todas las actividades, los estudiantes logran puntos y trofeos basados en su rendimiento. A medida que los estudiantes acumulan puntos en el iSTART avanzan desde el nivel 1 hasta el nivel 25, lo que implica el desbloqueo de juegos adicionales y de características personalizables, tales como nuevos colores de fondo del programa, o el estilo de sus avatares (ej. estilo de pelo, gafas, etc.). Finalmente los estudiantes son capaces de comprobar sus logros en el sistema (es decir, el número de veces que han completado cada actividad y los tipos de trofeos que han ganado) y la historia reciente respecto a la calidad de sus auto-explicaciones. A través de este tipo de características, los estudiantes pueden monitorear su desempeño.

En general, los estudiantes que interactúan tanto con el iSTART-2 como con sus versiones predecesoras han demostrado mejoras en su capacidad de auto-explicar y comprender textos científicos en comparación con los que no recibieron formación o tan sólo recibieron una breve instrucción (McNamara et al., 2004; McNamara, O'Reilly, Best, & Ozuru, 2006; McNamara, O'Reilly, Rowe, Boonthum, & Levinstein, 2007). Los estudiantes que usan las características basadas en el juego también reportan niveles más altos de motivación y de satisfacción en comparación con los estudiantes que no tienen acceso a estas características (Jackson & McNamara, 2013). Los trabajos recientes sobre el iSTART-2 se han focalizado en generar métodos para clasificar las interacciones de los estudiantes con el sistema y relacionar esas interacciones con el rendimiento. Este análisis resulta particularmente importante porque permite demostrar que el iSTART-2 ofrece a los estudiantes varias opciones dentro de la interfaz de práctica. No todas las opciones pueden ser igualmente beneficiosas para los estudiantes, lo que depende de los perfiles y diferencias individuales. Por ejemplo, un estudio reciente encontró que los estudiantes cuyos patrones de interacción fueron más ordenados también tuvieron un mejor desempeño en el sistema (Snow, Allen, Jacovina, & McNamara, 2015). El uso de este tipo de información permite

guiar la retroalimentación del programa, por ejemplo, proporcionando actividades sugeridas cuando las opciones de los estudiantes parecen ser desordenadas. Todo esto permite una mejor adaptación del iSTART-2 a las diferencias individuales, promoviendo un mejor rendimiento.

### *Proceso para la versión en español*

Para permitir que el iSTART-2 pueda ser aplicado a estudiantes de habla hispana, se está generando un trabajo de traducción y adecuación del programa al español (iSTART-E), un proceso que se inició en el año 2014. La traducción ha sido llevada a cabo a través de dos fases principales, traducción - programación de los videos, y traducción de los textos.

La fase de traducción de los videos ha implicado analizar los ocho videos en inglés incluidos en el programa y proponer ajustes para la versión en español tratando de mantener el sentido original de los mismos. Esto involucró hacer que las descripciones y explicaciones sobre las estrategias sean perfectamente entendibles para quienes hablan español, poniendo especial cuidado en los ejemplos que ofrece el sistema acerca de cada una de estas estrategias. Una vez que fue realizada la traducción de cada video ésta fue chequeada por un equipo de investigadoras argentinas dedicadas al tema de comprensión lectora con la intención de que ellas pudieran proponer ajustes, en términos de la comprensibilidad de los mismos, y de que el lenguaje empleado fuera lo más natural posible para el lector de habla hispana. Luego de esta revisión se integraron los comentarios en la versión definitiva de la traducción. Paralelamente se debió elegir la voz en español para el iSTART-E ya que en la fase de videos el agente educativo debe generar presentaciones orales y escritas acerca de cada estrategia. Para esto se optó entre las diversas alternativas de voces en español del programa “Loquendo Speech Packs” (de Character Builder/ Media Semantics), el que ofrece varias alternativas posibles. Para la decisión se solicitó el punto de vista de hablantes nativos

en español de diferentes nacionalidades (Chile, Argentina, España y México) para que manifestaran su opinión con respecto a cuál de las opciones de voz podría ser más neutral en términos de influencia dialectal. Considerando las diferentes opiniones se optó por “Carlos”, una de las voces del español norteamericano. Luego de incluir voces y textos en los videos se verificó que el discurso general del agente educativo funcionara bien en términos de la presentación de las lecciones, y sin errores en cuanto a la funcionalidad con el programa Character Builder. Se agregaron palabras y oraciones en español asociadas al *power point* del agente educativo, así como renovadas imágenes y textos que acompañan las lecciones presentadas. Luego de concluida esta fase se pidió la colaboración de hablantes en español (originarios de diferentes países de habla hispana) para que evaluaran qué tan comprensible y natural les resultaban los videos presentados. Producto de esta evaluación los participantes formularon algunos comentarios sobre expresiones, acentos u otros aspectos del lenguaje, y luego se les entrevistó acerca de esta práctica, de tal manera de obtener una idea lo más precisa posible de cada una de las observaciones. Como resultado de lo anterior se identificaron y registraron algunas expresiones que resultaron poco comunes para algunos de los participantes. En otros casos nos encontramos con palabras específicas como “automóvil” (*carro* en México, *coche* en España, *auto* en Chile) sobre las cuáles existen claras diferencias dialectales, y entonces debimos decidir por aquella expresión que fuera más comprensible y cercana para hablantes nativos en español de diferentes partes del mundo. De esta manera generamos la versión final para cada video del iSTART-E, las que fueron procesadas por el programa Character Builder a través de archivos en lenguaje xml. De esta manera obtuvimos los siguientes videos en español como parte de la primera fase de elaboración del iSTART-E: Un video de introducción que presenta el panorama general del programa; cinco videos con las cinco estrategias (parafraseo, puente, monitoreo, elaboración, predicción); un video con

un resumen de las lecciones y el último video de demostración que muestra cómo se pueden aplicar diversas estrategias para la comprensión de textos concretos.

La fase de práctica del programa iSTART-2 funciona a través del proceso de entrenamiento con alrededor de 32 textos y series de auto-explicaciones para cada uno de esos textos. La programación de este material -usando análisis semántico latente y palabras claves- permite que el programa pueda entregar una adecuada retroalimentación de las auto-explicaciones que generan los estudiantes acerca de los textos que leen. Por esta razón es que resulta fundamental contar con excelentes traducciones de los textos y de las respectivas auto-explicaciones. Este proceso se encuentra en una fase muy avanzada, y las traducciones de los textos se están revisando con el equipo de investigadoras argentinas para obtener una visión complementaria de las traducciones, de tal manera que se mantenga el ideal de una versión en español lo más neutral posible.

#### *Sobre el algoritmo de evaluación*

Como ya se ha mencionado previamente, la versión actual de iSTART 2 emplea un algoritmo de evaluación que mide la calidad global de las auto-explicaciones individuales (a nivel de frase). La asignación de estas puntuaciones a las auto-explicaciones posteriormente permite que el sistema iSTART proporcione a los estudiantes la retroalimentación relevante sobre cómo corregir sus estrategias a futuro (McNamara et al., 2007). Para calcular las puntuaciones de las auto-explicaciones, el algoritmo utiliza la información extraída de varias fuentes: la auto-explicación individual, la oración meta del texto que se auto-explica, y las frases del texto que rodean la oración meta. Estas fuentes de información son analizadas con medidas basadas en palabras y en análisis semántico latente (LSA; Landauer, McNamara, Dennis, & Kintsch, 2007).

Los índices basados en palabras en el algoritmo del iSTART proporcionan información acerca de las propiedades de *nivel inferior* de las auto-explicaciones. La información de nivel inferior se refiere a las características superficiales del texto, tales como la longitud del texto y el grado en que las palabras y frases se superponen a través de las oraciones. Estos índices de nivel inferior ayudan al sistema a identificar auto-explicaciones que son demasiado breves, muy similares a la oración meta, o irrelevantes para el texto. El LSA se utiliza para analizar las propiedades de *nivel superior* de las auto-explicaciones, proporcionando medidas sobre la similitud semántica entre dos palabras o secciones del texto. En términos generales, estos índices proporcionan información sobre el grado en que las auto-explicaciones de los estudiantes contienen información clave que se encontró en el texto e información importante que proviene desde fuera del texto (es decir, del conocimiento de mundo de los estudiantes). Cuando se combina con las medidas basadas en palabras, el índice de LSA proporciona evaluaciones más integrales sobre la calidad de las auto-explicaciones. El programa iSTART utiliza la información de los índices basándose en palabras y LSA con el fin de asignar a las auto-explicaciones una puntuación de 0 a 3. Una puntuación de "0" significa que una auto-explicación es demasiado corta para ser evaluada o contiene información irrelevante para el texto. Una puntuación de "1" indica que la auto-explicación se relaciona sólo con la información contenida en la oración meta. Las auto-explicaciones que reciben "2" integran información del texto que está más allá de la oración meta. Por último, el puntaje "3" se asigna a auto-explicaciones que incorporan información desde fuera del texto. Investigaciones previas indican que las puntuaciones del algoritmo de las auto-explicaciones muestran que el iSTART presenta una precisión similar a las evaluaciones realizadas por humanos (Jackson, Guess, & McNamara, 2010).



Para el programa iSTART-E, el algoritmo de evaluación del iSTART tendrá que someterse a una serie de cambios. En primer lugar, los índices basados en palabras tendrán que ser adaptados al español. Este proceso involucra primero traducción de las palabras y frases del inglés al español; a continuación, estas frases tendrán que ser validadas por hablantes nativos del español. Un segundo paso será encontrar o crear un gran corpus de textos en español con el que podremos construir un espacio de LSA en español, lo que nos permitirá calcular los índices de cohesión semántica. Por último, tendremos que evaluar la exactitud de este nuevo algoritmo en iSTART-E para lo cual será necesario recolectar un corpus de puntuaciones humanas de las auto-explicaciones en español. Sólo entonces, la precisión del algoritmo del iSTART-E se comparará con las puntuaciones asignadas por los evaluadores humanos, lo que nos permitirá determinar la eficacia del sistema.

### *Conclusiones*

En conclusión, el iSTART-E será una herramienta muy útil para entrenar las habilidades de comprensión lectora de los estudiantes de habla hispana. Este programa, como se ha señalado, proporciona lecciones y demostraciones en el uso de estrategias, además de un amplio proceso de entrenamiento utilizando diferentes tipos de juegos. Una vez concluida la fase de validación del piloto, podrán desarrollarse una serie de estudios experimentales que permitirán determinar cómo mejorar constantemente el impacto de esta herramienta teniendo en consideración factores tales como el tipo de retroalimentación, el impacto sobre la auto-regulación o las características metacognitivas de los sujetos. La práctica constante con esta herramienta sin duda ayudará a los estudiantes de habla hispana a funcionar más eficazmente en el uso de una de las habilidades necesarias para enfrentar el siglo XXI.

## Bibliografía

- Chi, M. T. H., de Leeuw, N., Chiu, M.-H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*, 439–477.
- Jackson, G.T., Guess, R.H., & McNamara, D.S. (2010). Assessing cognitively complex strategy use in an untrained domain. *Topics in Cognitive Science*, *2*, 127-137.
- Jackson, G. T., & McNamara, D. S. (2013). Motivation and performance in a game-based intelligent tutoring system. *Journal of Educational Psychology*, *105*, 1036-1049.
- Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (Eds.). (2007). *Handbook of Latent Semantic Analysis*. Mahwah, NJ: Lawrence Erlbaum.
- McNamara, D.S. (2004). SERT: Self-explanation reading training. *Discourse Processes*, *38*, 1-30.
- McNamara, D.S., Levinstein, I.B., & Boonthum, C. (2004). iSTART: Interactive strategy training for active reading and thinking. *Behavioral Research Methods, Instruments, & Computers*, *36*, 222-233.
- McNamara, D.S., O'Reilly, T., Best, R., & Ozuru, Y. (2006). Improving adolescent students' reading comprehension with iSTART. *Journal of Educational Computing Research*, *34*, 147-171.
- McNamara, D.S., O'Reilly, T., Rowe, M., Boonthum, C., & Levinstein, I.B. (2007). iSTART: A web-based tutor that teaches self-explanation and metacognitive reading strategies. In D.S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 397-420). Mahwah, NJ: Erlbaum.
- Snow, E. L., Allen, L. K., Jacovina, M. E., & McNamara, D. S. (2015). Does agency matter?: Exploring the impact of controlled behaviors within a game-based environment. *Computers & Education*, *26*, 378-392.