

UNIVERSIDAD DE COSTA RICA
SISTEMA DE ESTUDIOS DE POSGRADO

RESOLUCION DE DIAGRAMAS DE FLUJO Y PSEUDOCÓDIGO POR PARTE DE
ESTUDIANTES DE CIENCIAS DE LA COMPUTACIÓN DE LA UNIVERSIDAD DE
COSTA RICA.

Tesis sometida a la consideración de la Comisión del Programa de Estudios de Posgrado en
Ciencias Cognoscitivas para optar al grado y título de Maestría Académica en Ciencias
Cognoscitivas

LUIS ANDRÉS VILLALOBOS FERNÁNDEZ

Ciudad Universitaria Rodrigo Facio, Costa Rica

2020

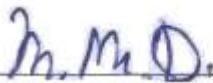
Dedicatoria

Este trabajo está dedicado a mi familia y a los docentes que me apoyaron a lo largo del proceso.

Agradecimientos

Al equipo asesor conformado por el Dr. Álvaro de Ossa, el Dr. Odir Rodríguez y el Dr. Adrián Vergara, además del profesor MSc. Armel Brizuela, quien colaboro en el uso del equipo de laboratorio y del software para el análisis estadístico.

Esta tesis fue aceptada por la Comisión del Programa de Estudios de Posgrado en Ciencias Cognoscitivas de la Universidad de Costa Rica, como requisito parcial para optar al grado y título de Maestría Académica en Ciencias Cognoscitivas



Dr. Mauricio Molina Delgado
Representante del decano



Dr. Álvaro de la Ossa Osegueda
Profesor Guía



Dr. Odir Rodríguez Villagra
Lector



Dr. Adrián Vergara Heidke
Lector



M.Sc. Armel Brizuela Rodríguez
Representante del director del programa de posgrado de Ciencias Cognoscitivas



Luis Andrés Villalobos Fernández
Sustentante

Tabla de contenidos

Dedicatoria y agradecimientos.....	ii
Hoja de aprobación.....	iii
Resumen.....	v
Abstract.....	vi
Listas de tablas y figuras.....	vii
Licencia de publicación.....	viii
Capítulo 1: Marco Referencial.....	1
Capítulo 2: Antecedentes.....	3
Capítulo 3: Marco Teórico.....	7
Capítulo 4: Metodología.....	13
Capítulo 5: Resultados.....	19
Capítulo 6: Conclusiones y recomendaciones.....	24
Referencias bibliográficas.....	26

Resumen

Se llevó a cabo un estudio de seguimiento ocular 20 estudiantes de ciencias de la computación de la Universidad de Costa Rica, la mitad de primer año y la otra mitad de cuarto año. Se les pidió que resolvieran 10 ejercicios de diagramas de flujo y 10 de pseudocódigo. En los expertos aproximadamente la mitad de las fijaciones se centran en las áreas de interés más relevantes para la resolución del problema, mientras que en los novatos aproximadamente un tercio de las fijaciones se concentran en esas áreas. Al calcular la densidad de fijaciones en las áreas de interés este patrón es mucho más evidente. En cuanto al rendimiento, se observaron diferencias significativas entre expertos y novatos, mostrando los primeros mejor rendimiento. En cuanto a la duración de la resolución de los ejercicios solo se encontraron diferencias significativas entre expertos y novatos al resolver pseudocódigos. En función de los resultados obtenidos se llevan a cabo recomendaciones para los y las profesores de los cursos de programación y para futuras investigaciones.

Palabras clave: ciencias cognitivas; diagramas; seguimiento ocular; programación; pseudocódigo.

Abstract

An eye tracking study was carried out on 20 students of computer science at the University of Costa Rica, half of the first year and the other half of the fourth year. They were asked to solve 10 flowchart and 10 pseudocode exercises. In the experts approximately half of the fixations focus on the most relevant areas of interest for the resolution of the problem, while in the rookies approximately one third of the fixations are concentrated in those areas. When calculating the density of fixations in the areas of interest this pattern is much more evident. Regarding performance, significant differences were observed between experts and novices, showing the first best performance. Regarding the duration of the resolution of the exercises, only significant differences were found between experts and novices when solving pseudocodes. Depending on the results obtained, recommendations are made for the teachers of the programming courses and for future research.

Key words: cognitive science; diagrams; eye-tracking; meta-analysis; programming; pseudocode.

Lista de tablas

Tabla 1.....	15
Tabla 2.....	20
Tabla 3.....	20
Tabla 4.....	21
Tabla 5.....	21
Tabla 6.....	22
Tabla 7.....	22

Lista de figuras

Figura 1.....	4
Figura 2.....	10
Figura 3.....	17
Figura 4.....	19
Figura 5.....	20



UNIVERSIDAD DE
COSTA RICA

SEP Sistema de
Estudios de Posgrado

Autorización para digitalización y comunicación pública de Trabajos Finales de Graduación del Sistema de Estudios de Posgrado en el Repositorio Institucional de la Universidad de Costa Rica.

Yo, Luis Andrés Villalobos Fernández con cédula de identidad 1-1367-0938, en mi condición de autor del TFG titulado Resolución de diagramas de flujo y Pseudo código por parte de Estudiantes de Ciencias de la Computación de la Universidad de Costa Rica

Autorizo a la Universidad de Costa Rica para digitalizar y hacer divulgación pública de forma gratuita de dicho TFG a través del Repositorio Institucional u otro medio electrónico, para ser puesto a disposición del público según lo que establezca el Sistema de Estudios de Posgrado. SI NO *

*En caso de la negativa favor indicar el tiempo de restricción: _____ año (s).

Este Trabajo Final de Graduación será publicado en formato PDF, o en el formato que en el momento se establezca, de tal forma que el acceso al mismo sea libre, con el fin de permitir la consulta e impresión, pero no su modificación.

Manifiesto que mi Trabajo Final de Graduación fue debidamente subido al sistema digital Kerwá y su contenido corresponde al documento original que sirvió para la obtención de mi título, y que su información no infringe ni violenta ningún derecho a terceros. El TFG además cuenta con el visto bueno de mi Director (a) de Tesis o Tutor (a) y cumplió con lo establecido en la revisión del Formato por parte del Sistema de Estudios de Posgrado.

INFORMACIÓN DEL ESTUDIANTE:

Nombre Completo: Luis Andrés Villalobos Fernández

Número de Carné: A77015 Número de cédula: 1-1367-0938

Correo Electrónico: lvillalobosfernandez@gmail.com

Fecha: 9/6/2020 Número de teléfono: 8466-2343

Nombre del Director (a) de Tesis o Tutor (a): Dr. Álvaro de la Ossa Osegueda

FIRMA ESTUDIANTE

Nota: El presente documento constituye una declaración jurada, cuyos alcances aseguran a la Universidad, que su contenido sea tomado como cierto. Su importancia radica en que permite abreviar procedimientos administrativos, y al mismo tiempo genera una responsabilidad legal para que quien declare contrario a la verdad de lo que manifiesta, puede como consecuencia, enfrentar un proceso penal por delito de perjurio, tipificado en el artículo 318 de nuestro Código Penal. Lo anterior implica que el estudiante se vea forzado a realizar su mayor esfuerzo para que no sólo incluya información veraz en la Licencia de Publicación, sino que también realice diligentemente la gestión de subir el documento correcto en la plataforma digital Kerwá.

Capítulo 1: Marco referencial

Justificación

La carrera de Ciencias de la Computación tiene la particularidad de que el contenido de los textos académicos es representado no solo en el lenguaje natural, sino en gran parte en representaciones artificiales. Como recurso didáctico se emplea gran cantidad de medios visuales como diagramas, gráficos, infografías y simulaciones, entre otros. Esto hace necesario llevar a cabo investigaciones y elaborar propuestas pedagógicas, referentes a la comprensión de la materia, diferenciadas de otras carreras (Rueda, 2013).

Un posible enfoque del problema pedagógico en las Ciencias de la Computación es el análisis de las estrategias con las que los estudiantes observan las representaciones artificiales en los textos de estudio para comprenderlas, asociado a su desempeño en la comprensión de esas representaciones. Una técnica disponible para este fin es el rastreo ocular o eye tracking.

El eye tracking es una técnica que permite registrar los movimientos oculares cuando se está llevando a cabo alguna actividad. Estos movimientos se relacionan con el nivel de atención que el sujeto observado les dedica a los elementos de la tarea que está resolviendo. Por esta razón, su uso ha aumentado considerablemente en la investigación pedagógica.

Objetivo general

- Describir las diferencias entre novatos y expertos en la resolución de ejercicios de programación planteados como diagrama de flujo o pseudocódigo.

Objetivos específicos

- Estimar las diferencias entre novatos y expertos en los tiempos de resolución de ejercicios de programación de diagramas de flujo y pseudocódigo.
- Estimar las diferencias entre novatos y expertos en el número de ejercicios de programación de diagramas de flujo y pseudocódigo resueltos correctamente.
- Explorar las diferencias entre novatos y expertos en la trayectoria de los movimientos oculares mientras intentan resolver ejercicios de programación de diagramas de flujo y pseudocódigo.

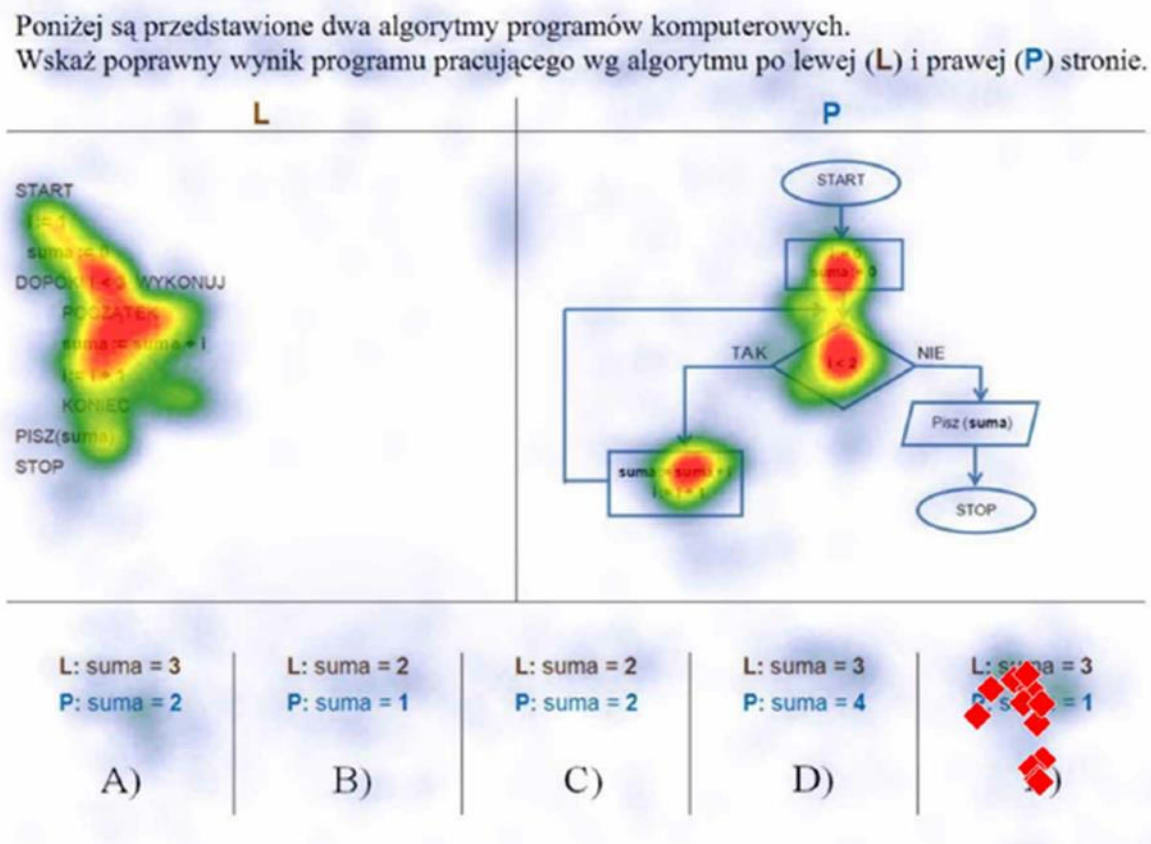
Palabras clave

- Ciencias cognitivas.
- Diagramas de flujo.
- Seguimiento ocular.
- Programación.
- Pseudocódigo.

Capítulo 2: Antecedentes

El único antecedente que se ha encontrado sobre el estudio del proceso de comprensión de diagramas de flujo, ayudado con el uso de eye tracking, es un estudio llevado a cabo por (Andrzejewska, Stolinska, Blasiak, Peczkowski, Rosiek, Rozek, 2015). En esa investigación, la tecnología de seguimiento ocular se utilizó para seguir el proceso de resolver problemas algorítmicos. Estos eran presentados en dos variantes comparables: un pseudocódigo y un diagrama de flujo. Los datos de rastreo de los ojos dieron como resultado un conjunto de medidas cualitativas (películas que registran las trayectorias de la mirada) y cuantitativas, que permitieron detectar e interpretar las diferencias en las estrategias de resolución de tareas entre los que encontraron la respuesta correcta y los que no. Los resultados confirman una hipótesis de que el uso de la notación formal característica de un lenguaje de programación para presentar algoritmos es a menudo una dificultad práctica en el proceso de resolver incluso tareas simples. Este estudio abre una nueva dirección de la investigación: muestra que la tecnología de rastreo ocular puede utilizarse para optimizar el proceso de aprendizaje de la programación.

Figura 1: Registro de movimientos oculares en diagramas de flujo.



Fuente: Andrzejewska *et al* (2015).

En cuanto a los patrones de movimiento ocular, se encontró que los estudiantes se centraban sobre todo en los símbolos que representan procesos y los que representan decisión, en el trabajo no se hace énfasis en el rol jugado por la visión periférica (Andrzejewska *et al*, 2015).

El curso original siguió la estructura convencional para un curso de base de datos, cubriendo el diseño de la base de datos primero, luego el desarrollo de la base de datos. El análisis mostro que el contenido del curso convencional era apropiado, pero los materiales didácticos utilizados eran demasiado complejos, especialmente para los estudiantes principiantes.

El rediseño de los materiales de instrucción aplico la teoría de la carga cognitiva para eliminar la atención dividida y los efectos de redundancia, para proporcionar ejemplos y subobjetivos adecuados, e incluyo una re-secuencia extensa del contenido.

El enfoque se dirigió principalmente a los estudiantes de medio a bajo rendimiento y los resultados mostraron una mejora significativa para la última cohorte, con una reducción del 34% en la tasa de fracaso del examen, después del rediseño del curso (Mason, Seton & Cooper, 2016)

Cardie (1999) realizó un estudio en la Universidad de Cornell en Nueva York, tenía como objetivo mostrar que las limitaciones psicológicas en el procesamiento de la información humana pueden ser utilizado con eficacia para guiar la selección de función para establecer el aprendizaje basado en casos de conocimientos lingüísticos. Teniendo en cuenta como entrada una representación del caso para una tarea de aprendizaje del lenguaje natural, el algoritmo selecciona los sesgos cognitivos relevantes para la tarea y luego automáticamente modifica la representación en respuesta a los sesgos cognitivos.

Las variables cognitivas que se tomaron en cuenta fueron. 1) accesibilidad del tema, se refiere a en qué medida el tema que se está tratando es de conocimiento del público al que se dirige, por ejemplo, que sea de un nivel muy elevado para su nivel educativo o de otro campo de conocimiento diferente. 2) Restricción de memoria, alude a la cantidad de caracteres que posee el texto a analizarse, de forma que implica mayor carga de memoria. 3) Tiempo de exposición, es la cantidad de tiempo que el sujeto se le otorga para poder ejecutar la tarea de comprensión lectora (Cardie, 1999).

Por medio de validaciones cruzadas, que es una técnica para evaluar los resultados de un análisis estadístico y garantizar que son independientes de la partición de datos, para estimar como de preciso es un modelo. La restricción de memoria y la accesibilidad al tema fueron las variables que más influyeron en los resultados de los sujetos evaluados (Cardie, 1999).

Según, Elousa, García, Vila, Gómez y Gil (2013) la investigación sobre la intervención de comprensión de lectura puede integrarse en clases completas, en pequeños grupos o en instrucción individual. Las intervenciones en la comprensión de lectura se pueden implementar en la escuela desarrollando estrategias de comprensión específicas tales como enseñar la idea principal y habilidades de resumen. Sin embargo, algunos estudios recientes han comenzado a desarrollar un enfoque diferente que se centra en los procesos ejecutivos de memoria de trabajo.

El papel del sistema de memoria de trabajo en la comprensión de lectura ha sido destacado en los modelos clásicos de comprensión del texto. Comprensión de lectura es a menudo un

proceso complejo que se logra dentro de la memoria de trabajo y requiere una gran cantidad de recursos cognitivos, incluyendo la eficacia de almacenamiento y procesamiento. La comprensión lectora es, pues, una tarea en la que los procesos ejecutivos están altamente implicados. El programa específico de capacitación enfocado a las funciones ejecutivas de la memoria de trabajo es un ejemplo de lo que los investigadores podrían hacer para comprender cómo, cuándo y por qué el proceso de comprensión de la lectura puede mejorarse en el entorno escolar. En este trabajo los investigadores presentaron un programa de capacitación basado en el trabajo Los procesos ejecutivos de repetición de la memoria, la retroalimentación y el ajuste gradual de la dificultad y uno que demuestra la efectividad de la capacitación adaptativa, como algunos estudios recientes han sugerido (Elosua et al, 2013). Yi, Yu & Ying (2014) llevaron a cabo un meta-análisis, en el que se revisaron 17 estudios de instrucción de estrategia metacognitiva sobre la comprensión lectora de los estudiantes en contextos de lectura computarizados. En general, algunos casos de instrucción de estrategia metacognitiva tienden a ser más eficaces que otros. Además, los efectos de la instrucción parecían variar según las características de los participantes.

Al revisar los resultados de los 17 estudios cuantitativos, sin embargo, los efectos de estos diferentes ejemplos de instrucción de estrategia metacognitiva siguen siendo mixtos. En general, se pueden extraer algunas tendencias interesantes de este metanálisis: 1) la enseñanza de la estrategia metacognitiva parece ser más eficaz para ayudar a los estudiantes a comprender los textos científicos, 2) la regulación se considera como una forma más eficaz de la instrucción, 3) los estudiantes que recibieron La instrucción no necesariamente superaron a sus homólogos, y 4) las habilidades de lectura de los estudiantes jugaron un papel interesante (Yi, Yu & Ying, 2014).

Capítulo 3: Marco teórico

Eye tracking

El cerebro procesa la entrada sensorial concentrándose en componentes específicos de todo el reino sensorial para que los lugares de interés, sonidos, olores y similares puedan examinarse con mayor atención al detalle que los estímulos periféricos. Esto es particularmente cierto de la visión. La inspección de la escena visual se realiza un elemento a la vez, no en totalidad. Es decir, la visión humana es un proceso fragmentario que depende de la integración perceptual de las pequeñas regiones para construir una representación coherente del todo (Duchowski, 2007).

Pueden hacerse dos observaciones significativas relevantes para el análisis del movimiento ocular. En primer lugar, basándose en la funcionalidad de los movimientos oculares, sólo se deben modelar tres tipos de movimientos para comprender mejor la localización manifiesta de la atención visual. Estos tipos de movimientos oculares son fijaciones, búsquedas lisas y sacadas. En segundo lugar, basándose en las características de la señal y los circuitos neuronales subyacentes plausibles, los tres tipos de movimientos oculares pueden ser aproximados por un sistema Linear, Time-Invariant (LTI). Es decir, un filtro lineal, para ejemplos de filtros lineales aplicables a la detección de sacadas (Duchowski, 2007).

El requisito primario del análisis del movimiento ocular, en el contexto del diseño del sistema contingente de la mirada, es la identificación de fijaciones, sacadas y persecuciones lisas. Se supone que estos movimientos proporcionan evidencia de atención visual voluntaria y manifiesta. Esta suposición no excluye la utilidad involuntaria plausible de estos movimientos, o inversamente, el uso encubierto de estos movimientos oculares (por ejemplo, como en el caso de la atención parafoveal). Las fijaciones corresponden naturalmente al deseo de mantener la mirada en un objeto de interés. Del mismo modo, las actividades se utilizan de la misma manera para los objetos en movimiento suave. Los movimientos sacádicos se consideran manifestaciones del deseo de cambiar voluntariamente el foco de atención (Duchowski, 2007).

El tiempo que un usuario fija un área en particular en la pantalla nos ayuda a entender si él / ella está prestando especial atención a un elemento visual específico. La duración de las fijaciones es a menudo extremadamente corta y se representa típicamente en milisegundos.

Desafortunadamente, la duración es una medida particularmente difícil interpretar. Hay muchas razones posibles por las que un usuario fija durante un período de tiempo relativamente corto o largo en un área particular. ¿Están confundidos por este elemento? ¿Lo encuentran atractivo? Para entender estas medidas necesitamos mirar los datos en contexto con otros métodos de investigación o con tareas cognitivas que permitan descartar la mayor cantidad de explicaciones rivales a la hipótesis de partida del estudio (Romano & Schall, 2014).

A continuación, se muestran algunos conceptos teóricos relativos a los efectos de los estímulos visuales en la comprensión.

Principio de contigüidad espacial

Este principio afirma que las personas aprenden mejor cuando las palabras e imágenes relacionadas se presentan más próximas unas de otras (Ware, 2008).

En la misma línea de este, encontramos el Principio de Proximidad de la Gestalt, el cual postula que aquellos objetos que aparecen próximos, se perciben como relacionados o pertenecientes a una misma categoría o grupo.

Tiene mayor aplicación cuando:

- El receptor no está familiarizado con el material, es decir, es novato en la materia.
- El diagrama no se entiende por sí solo. Debido a que hace referencia a conceptos abstractos o que no son visualizados cotidianamente. Por ejemplo, conceptos matemáticos o de biología celular.
- El material mostrado es complejo. Dado que implica gran cantidad de información, la cual es poco intuitiva.

En este sentido, los diagramas de flujo podrían ser una herramienta de utilidad para personas que estén en las etapas iniciales de su proceso de aprendizaje de las representaciones de algoritmos, aunque no necesariamente para los expertos.

Teoría de la codificación dual

De acuerdo con esta teoría, la información verbal y no verbal son procesadas por diferentes sub sistemas cognitivos. Por lo tanto, se da un incremento en la retención de esos elementos cuando se presentan juntos (Paivio, 1986).

- Códigos analógicos: la relación entre la representación y significado es motivado naturalmente (imágenes).
- Códigos digitales: la relación entre la representación y significado es arbitraria (letras y números).

Saliencia




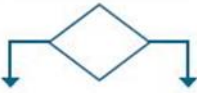
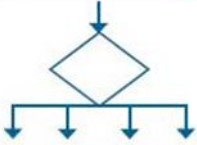



Dada la gran afluencia de información visual que nos rodea, es indiscutible la existencia de un mecanismo selectivo que nos permite descartar la información que no resulta útil para centrarnos en la que necesitamos para alcanzar los objetivos concretos que nos proponemos. A pesar del buen funcionamiento de este mecanismo selectivo, hay ocasiones en las que nos sorprendemos atendiendo a información irrelevante. Esto se debe a procesos de naturaleza “bottom-up” que resultan involuntarios al estar gobernados, no por nuestros intereses, sino por las propiedades inherentes de los estímulos, las cuales parecen capturar nuestra atención. Con esto, se pone de manifiesto que los mecanismos que gobiernan los procesos selectivos de la atención pueden estar determinados por dos tipos diferentes de información que controlan el despliegue atencional: (a) por las metas, intenciones y objetivos que se propone el sujeto –información top-down- y (b) por la propia información que contienen los estímulos –información bottom-up (Sampedro, Blanco, Ponte, Leiros, 2004).

En concreto, Bacon y Egeth (1994), defienden que la captura atencional está determinada por la saliencia de los estímulos (componente bottom-up) únicamente cuando los observadores utilizan un “modo de detección tipo singleton“, es decir buscan por la ocurrencia del elemento que difiere del resto.

Diagramas de flujo

Un diagrama de flujo es una representación pictórica que representa el flujo de pasos de un algoritmo computacional (ver figura 2). Los elementos se representan mediante iconos simples, como círculos, rectángulos, diamantes u otras formas, con líneas y flechas que representan las conexiones entre los eventos y la dirección o el orden en que se producen. Un diagrama de flujo indica secuencias y puntos de decisión, así como puntos de inicio y de parada (Tsapatsoulis, 2014).

Figura 2: Simbología de los diagramas de flujo.

Símbolo	Significado
	Terminal /Inicio.
	Entrada de datos.
	Proceso.
	Decisión.
	Decisión múltiple.
	Imprimir resultados.
	Flujo de datos.
	Conectores.

Fuente: Pinales, F y Velázquez, C (2014). Página 18.

La simbología mostrada anteriormente es la oficializada por el Instituto Nacional de Normalización Estadounidense (ANSI). Existiendo otros formatos, como el de la Sociedad Americana de Ingenieros Mecánicos (ASME) o el Instituto Alemán de Normalización (DIN).

La capacidad de construir, leer e interpretar diagramas de flujo es una habilidad necesaria para todos los tecnólogos. Los símbolos de diagramas de flujo se reconocen universalmente a través de múltiples disciplinas y permitirán a un profesional del área comunicar gran cantidad de información de forma precisa, eficiente y sucinta. Además, programadores y diseñadores de software y multimedia usan rutinariamente diagramas de flujo para ilustrar las capacidades de software, el rendimiento requerido del usuario y la respuesta de la computadora. El alto nivel de interactividad y ramificación involucrados en el desarrollo de software son justificación suficiente para que los informáticos deban dominar esta forma de representación de la información (Tsapatsoulis, 2014).

No hay un conjunto rígido de reglas para construir diagramas de flujo, pero los siguientes pasos sirven como directrices para garantizar la coherencia y precisión de los diagramas a elaborarse (Tsapatsoulis, 2014):

1. Describir el proceso que se va a llevar a cabo de forma muy concisa.
2. Comenzar por el evento que inicia el proceso.
3. Agregar cada acción sucesiva en forma concisa. Evitando descripciones ambiguas.
4. Cuando el organigrama se ramifica en numerosas opciones, empezar por la más importante o más probable y después desarrolle las demás.
5. Continuar describiendo cada evento, acción o decisión que se produce en secuencia, hasta llegar a la conclusión.

Evaluación de la memoria de trabajo por medio del eye tracking

A continuación, se procede a describir los elementos de interés a tomar en cuenta al llevar a cabo el análisis de la carga cognitiva por medio del eye tracking (Zagermann, Pfeil, Reiterer, 2016).

Fijaciones

Indica el número de veces que el usuario miro a una cierta área de interés. La duración de la fijación se ha relacionado con el nivel de carga cognitiva. Con una alta duración de fijación indicando un alto grado de carga cognitiva. A su vez se da una disminución en la tasa de fijación.

Movimientos sacádicos

La disminución en la velocidad de la sacada es indicativa de cansancio por parte de la persona, mientras que un aumento de esta se relaciona con un alto grado de dificultad de la tarea que se está llevando a cabo. Esto se refleja en sacadas más largas y más rápidas, conforme aumenta la dificultad del ejercicio.

Capítulo 4: Metodología

4.1. Participantes

Para este trabajo se reclutó a 20 estudiantes de la carrera de ciencias de la computación de la Universidad de Costa Rica. Todas las personas participantes firmaron un consentimiento informado, el cual fue aprobado por el comité ético-científico de la Universidad de Costa Rica. De ellos 10 fueron caracterizados como novatos debido a que apenas estaban cursando el primer semestre de la carrera, y por lo cual apenas estaban aprendiendo a programar, se excluyó del estudio a estudiantes que ya hayan llevado algún tipo de aprendizaje formal en programación antes de ingresar a la universidad y 10 expertos, los cuales estaban cursando el segundo semestre del cuarto año de carrera, razón por la cual ya han aprobado todos los cursos de programación de la carrera. En ambos grupos tuvo que descartarse un participante masculino debido a errores en la obtención de los datos, debido a movimientos de la cabeza durante la prueba. Por lo que el total de participantes fue 18 personas, 9 novatos y 9 expertos. La muestra es mayoritariamente masculina, en el grupo de novatos participaron 7 hombres y 2 mujeres, mientras que en el de expertos 8 hombres y 1 mujer. Todas las personas participantes eran mayores de edad, los novatos tenían una edad promedio de 18,5 años, mientras que los expertos 22,3 años.

4.2. Hardware y Software

Las personas participantes se sentaron confortablemente en una silla capaz de ajustarse a la estatura de cada persona, con sus cabezas estabilizadas por un soporte. Esto para eliminar las distracciones periféricas. Los participantes fueron evaluados individualmente y los movimientos oculares se registraron mediante un dispositivo de rastreo ocular *EyeLink 1000 Plus*, con una frecuencia de muestreo de 1000 Hz y un seguimiento binocular. Los ejercicios se mostraron en un monitor *BenQ XL2430T* de 24 pulgadas (53 cm de ancho y 30 cm de alto), cuya resolución es de 1920 x 1080 píxeles y cuya frecuencia de actualización es de 60 Hz. Este controlado por una computadora *Mac mini*, con un procesador *Intel Core i5* de 2.6 GHz. Los participantes se sentaron a 55 cm de la pantalla de la computadora. La calibración realizada fue de 9 puntos, registrándose solo la actividad del ojo derecho, las personas que presentaban problemas de la vista utilizaron sus anteojos durante la prueba. Para responder

las personas emplearon un teclado, en el que debían apretar las teclas 1, 2, 3 o 4, según la opción que eligieran en cada pregunta, el programa automáticamente desplegaba una pantalla de calibración, en la cual la persona debía dirigir su mirada hacia el centro de la pantalla, una vez cumplida el drift check, automáticamente se desplegaba el siguiente ítem.

4.3. Estímulos y procedimiento

El software en el que se programó el experimento es el Experiment Builder. El cual permite controlar todos los parámetros del experimento. Para este en concreto se trabajó con 20 estímulos, 10 de los cuales son de diagramas de flujo y 10 pseudocódigos, a su vez en estos grupos de ítems la mitad son de identificación de errores y la otra mitad de identificar la función del algoritmo. Los estímulos se presentaron en orden aleatorio y los participantes tuvieron un tiempo límite de 3 minutos para resolver los ítems.

El tipo y tamaño de letra empleado es la times new roman 26, todos los ítems seguían la misma estructura: en la esquina superior izquierda se ubican las instrucciones y las opciones de respuesta, en la esquina inferior izquierda se ubicaba la caja en que se explican las siglas de las variables del problema, y en el lado derecho el problema expresado como pseudocódigo o diagrama de flujo. Independientemente del ejercicio, estos siguen la misma estructura, el comando de inicio, una serie de comandos de acción, la decisión y la impresión y el fin (ver figura 1).

4.4 Diseño

Como covariables se empleó la inteligencia fluida, medida por medio del test de factor G de Cattell, la velocidad de procesamiento medida por medio de la prueba de dígitos y símbolos y la memoria de trabajo mediante la prueba de dígitos en orden inverso y en orden directo.

Tabla 1

Puntuaciones de las pruebas aplicadas como covariables

Directa Cattell	Media	T	DF	P-Value
Expertos	24.9	0.22	15.26	0.823
Novatos	24.7			
Centil Cattell	Media	T	DF	P-Value
Expertos	82.8	0.52	14.11	0.606
Novatos	80.6			
CI Cattell	Media	T	DF	P-Value
Expertos	108	0.22	15.26	0.823
Novatos	107			
Dígitos y símbolos	Media	T	DF	P-Value
Expertos	60.1	0.26	15.94	0.797
Novatos	59.2			
Dígitos directos	Media	T	DF	P-Value
Expertos	9.56	2.13	13.50	0.051
Novatos	9.11			
Dígitos inversos	Media	T	DF	P-Value
Expertos	8.67	2	15.75	0.063
Novatos	8.22			

Tabla 1. Muestra las puntuaciones obtenidos en las pruebas aplicadas como covariables, para determinar si las diferencias observadas entre ambos grupos se debieron a su nivel de experticia en la materia o a un mayor nivel de inteligencia fluida, memoria de trabajo y velocidad de procesamiento. En todas las puntuaciones el P-Value es superior a 0.05.

Los novatos obtuvieron una puntuación CI Cattell de 106,66, mientras que los expertos 107,55 en dicha prueba. Mientras que en la prueba de dígitos y símbolos los novatos obtuvieron una puntuación directa de 59,22 y los expertos 60,11; por lo que no se aprecian diferencias significativas entre ambas poblaciones con respecto a sus capacidades cognitivas, por lo que la principal diferencia entre ambos grupos es su nivel de experiencia con respecto a la materia.

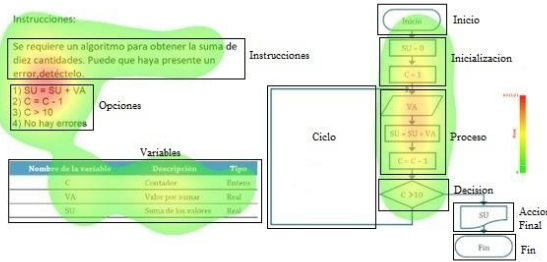
Las áreas de interés fueron definidas en base al criterio experto de profesores de la Escuela de Ciencias de la Computación de la Universidad de Costa Rica. Estas fueron definidas en función de los procesos que se llevan a cabo al ejecutar un algoritmo. El inicio marca el principio del proceso, la inicialización agrupa las primeras acciones, el proceso la entrada y acciones posteriores, por aparte está la decisión, el ciclo, la acción final y el fin.

El criterio experto predecía que las áreas más importantes para la resolución del problema son el proceso, la decisión y la inicialización. Mientras que las opciones, las variables y las instrucciones lo son para informarse de las características del problema.

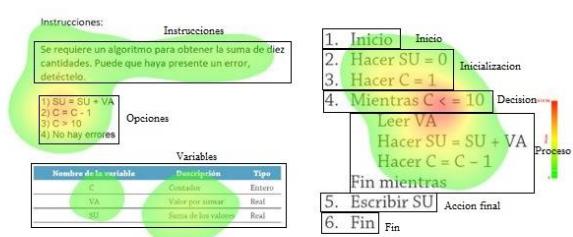
En la figura 1 se pueden apreciar las distintas áreas de interés, así como el patrón de visualización promediado en dos ejercicios, diagrama de flujo en el lado izquierdo y pseudocódigo en el lado derecho, novatos en la parte superior y expertos en la inferior.

Figura 3
Áreas de interés y fijaciones

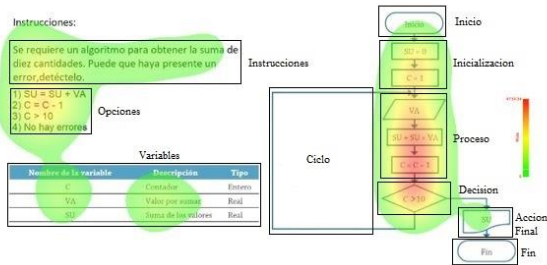
Ítem 1 novatos



Ítem 11 novatos



Ítem 1 expertos



Ítem 11 expertos



Figura 3. En esta imagen se muestra uno de los ejercicios que se le presento a los participantes, en formato de diagrama de flujo (izquierda) y pseudocódigo (derecha). El patrón de visualización de los novatos (arriba) y expertos (abajo), muestra diferencias en cuanto a las áreas en las que se dio una mayor cantidad de fijaciones.

4.5. Análisis estadístico

Los datos se extrajeron del programa Data Viewer (SR Research, 2017). Dicho programa genera una base de datos en la cual para cada una de las áreas de interés de cada ejercicio, en cada una de las personas participantes, se genera una fila de datos, obteniéndose 3286 observaciones (18 participantes, 20 ejercicios, 9 – 10 áreas de interés por ejercicio, el área de interés ciclo no se encuentra presente en los pseudocódigos). Para cada una de esas observaciones se dispone de las fijaciones totales en dicha área, la duración en milisegundos de las fijaciones en dicha área, la respuesta dada por el usuario en cada ejercicio y la apertura máxima de la pupila, las puntuaciones de dichas variables se promediaron para llevar a cabo los análisis. El análisis estadístico de los datos se llevó a cabo en el programa R Studio (Wickham & Grolemund, 2017). Se seleccionó este programa estadístico debido a que permite la instalación de varios paquetes complementarios que facilitan el proceso de análisis, en concreto se utilizaron, ggplot2 (Wickham, Chang, Henry, Pedersen, Takahashi, Wilke, Woo, 2020), dbplyr (Wickham, Francois, Henry, Muller, 2020), naniar (Tierney, Cook, McBain, Fay, 2020) y Rmarkdown (Allaire, Xie, Mc Pherson, Luraschi, Ushey, Atkins, Wickham, 2020). Se llevaron a cabo gráficos de barras con error estándar y pruebas T.

Capítulo 5: Resultados

En la figura 4 se muestra que, en los diagramas de flujo, los expertos presentan un mayor número de fijaciones en las áreas de decisión, inicialización y proceso, tomando en cuenta el error estándar, mientras que los novatos presentan un mayor número de fijaciones en las áreas de ciclo, opciones y variables, tomando en cuenta el error estándar.

Figura 4

Gráficos de barras con error estándar diagramas de flujo expertos - novatos

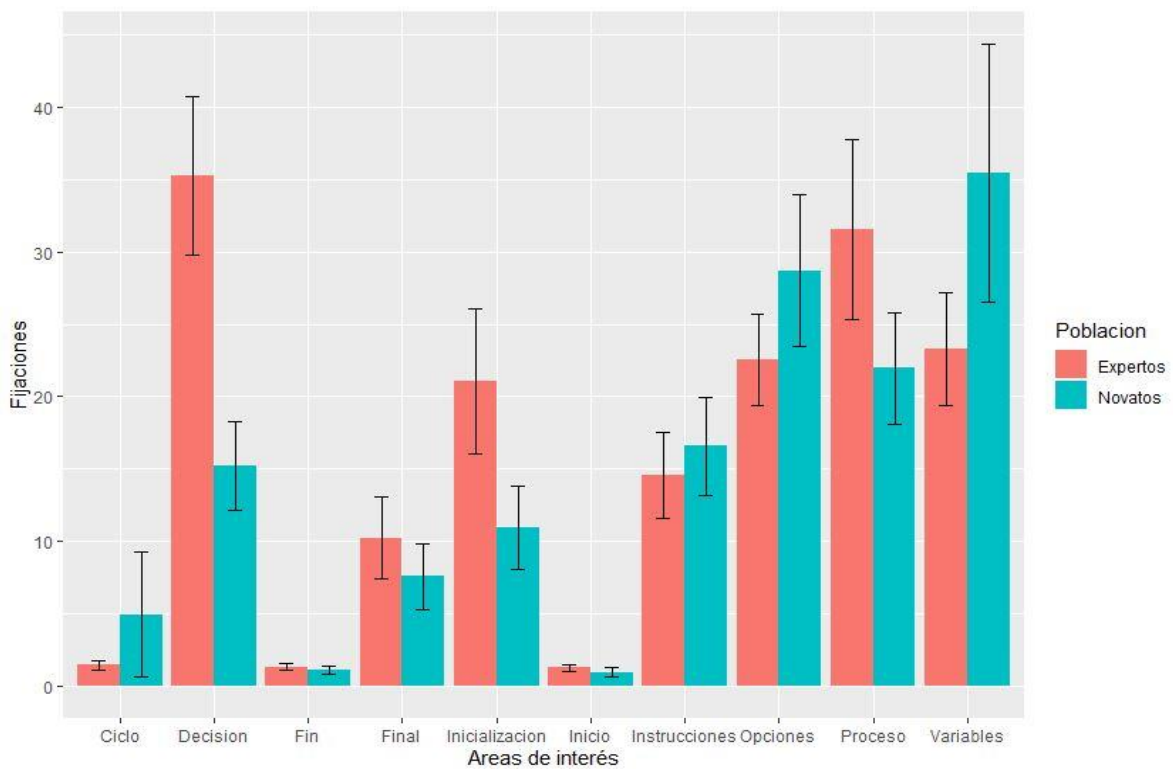


Figura 4. Se muestran diferencias entre expertos y novatos en el número de fijaciones en las diversas áreas de interés de los diagramas de flujo.

Como se puede apreciar en la figura 5, en los pseudocódigos, los expertos presentan un mayor número de fijaciones en el área de decisión, tomando en cuenta el error estándar, mientras que los novatos presentan un mayor número de fijaciones en las áreas de instrucciones, opciones y variables. Un mayor número de fijaciones puede ser un indicativo de mayor tiempo de procesamiento.

Figura 5

Gráficos de barras con error estándar pseudocódigos expertos - novatos

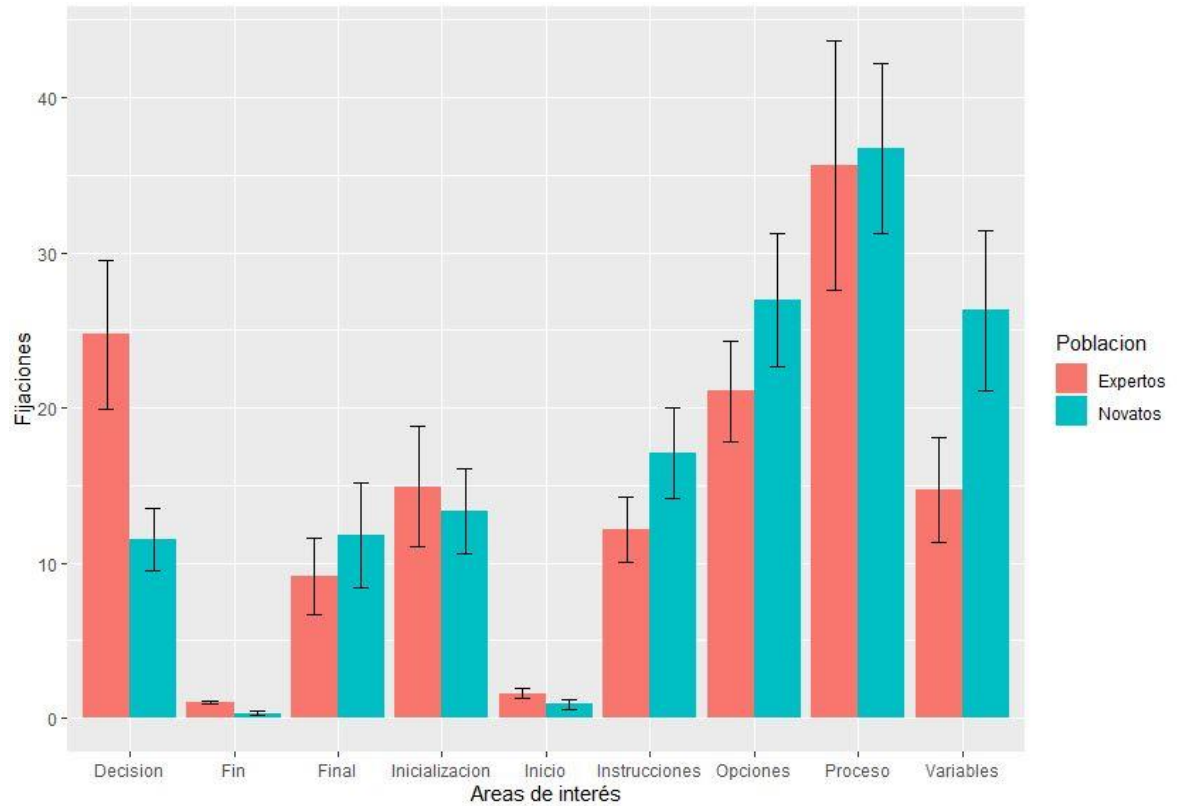


Figura 5. Se muestran diferencias entre expertos y novatos en el número de fijaciones en las diversas áreas de interés de los pseudocódigos.

Tabla 2

Prueba T Student diferencias en rendimiento entre expertos y novatos en diagramas de flujo

Población	N	Media	T	DF	P-Value
Expertos	9	0.876	10.833	1507.5	2.2e-16
Novatos	9	0.661			

Tabla 3

Prueba T Student diferencias en rendimiento entre expertos y novatos en pseudocódigos

Población	N	Media	T	DF	P-Value
Expertos	9	0.791	4.100	1578.3	4.337e-05
Novatos	9	0.702			

Las pruebas de t de student llevadas a cabo para determinar si hay diferencias significativas entre novatos y expertos en cuanto al rendimiento, mostro que, en el caso de los diagramas de flujo, el p-value es de $2.2e-16$, dado que es < 0.05 , se rechaza la hipótesis nula y podemos afirmar que las muestras difieren en su media (tabla 2). Con respecto a los pseudocódigos, el p-value es de $4.337e-05$, dado que es < 0.05 , se rechaza la hipótesis nula y podemos afirmar que las muestras difieren en su media (tabla 3).

Tabla 4

Prueba T Student diferencias duración en milisegundos entre expertos y novatos en diagramas de flujo

Población	N	Media	T	DF	P-Value
Expertos	9	47880.46	1.497	1629.3	0.1344
Novatos	9	45626.37			

Tabla 5

Prueba T Student diferencias duración en milisegundos entre expertos y novatos en pseudocódigos

Población	N	Media	T	DF	P-Value
Expertos	9	36544.13	-5.432	1509.1	6.485e-08
Novatos	9	43144.68			

Las pruebas de t de student llevadas a cabo para determinar si hay diferencias significativas entre novatos y expertos en cuanto al tiempo que duraban resolviendo los problemas, mostro que, en el caso de los diagramas de flujo, el p-value es de 0.1344, dado que es > 0.05 , no se rechaza la hipótesis nula y podemos afirmar que las muestras no difieren en su media (tabla 4). Con respecto a los pseudocódigos, el p-value es de $6.485e-08$, dado que es < 0.05 , se rechaza la hipótesis nula y podemos afirmar que las muestras difieren en su media (tabla 5).

Tabla 6

Prueba T Student diferencias duración en milisegundos entre expertos y novatos en diagramas de flujo, excluyendo la primera fijación.

Población	N	Media	T	DF	P-Value
Expertos	9	40934.92	3.06	1280.8	0.002
Novatos	9	36022.35			

Tabla 7

Prueba T Student diferencias duración en milisegundos entre expertos y novatos en pseudocódigos, excluyendo la primera fijación.

Población	N	Media	T	DF	P-Value
Expertos	9	30610.44	-3.14	1120.5	0.001
Novatos	9	34811.36			

Como medida de control se volvió a calcular las diferencias en duración, eliminando la primera fijación, las diferencias entre ambas poblaciones siguen siendo significativas.

Al tomarse en cuenta como variables de control la inteligencia fluida y la memoria de trabajo de las personas participantes, se descarta que las diferencias observadas se deban a diferencias en las capacidades cognitivas de los participantes. Esto es de relevancia para futuras investigaciones con seguimiento ocular, ya que debido a lo especializado que es el equipo y el tiempo de aplicación (en torno a una hora por participante, incluyendo firma de consentimiento informado y otras pruebas a aplicar) las muestras en este tipo de estudios suelen ser pequeñas, entorno a los 15 – 30 participantes (Gegenfurther, Lehtinen, Saljo, 2011). Por lo que contar con métodos adicionales para asegurar la confiabilidad de los datos es importante.

Al disponerse de una batería de ítems más extensa que en investigaciones anteriores, se asegura una mayor confiabilidad externa de los datos, ya que se simula con mayor fiabilidad el tipo de experiencia a la que se verán sometidos los estudiantes en el entorno académico.

Al mostrarse que hay un patrón diferenciado de visualización entre novatos y expertos, mostrándose que los segundos tienen un mayor rendimiento. Se debe procurar que los

novatos adopten en su formación estrategias similares para poder resolver los problemas de representación de algoritmos.

En concreto, se identificó que los novatos llevan a cabo una mayor cantidad de fijaciones en áreas de interés de tipo informativo, mientras que los expertos se centran en las áreas que contienen la información relevante para la resolución del problema.

Entre las recomendaciones que se podrían llevar a cabo para mejorar la enseñanza de la representación de algoritmos en estudiantes de primer año, están:

1. Presentar ejemplos en que las áreas más importantes estén representadas de otro color. Pudiéndose utilizar este sistema tanto en los pseudocódigos como en los diagramas de flujo.
2. Variar el sistema de representación de los diagramas de flujo, de forma que las acciones, las cuales normalmente se representan como un rectángulo, independientemente de su ubicación en el problema, tengan una representación diferente, en función de si forman parte del proceso de inicialización o proceso.
3. En los diagramas de flujo usados en esta investigación, la línea que señala el ciclo está bastante del resto del diagrama, lo cual pudo influir en las pocas fijaciones en esa área. Por lo que se podría utilizar diagramas en los que dicha línea está más cercana, para facilitar su identificación.

Capítulo 6: Conclusiones y recomendaciones

En los expertos aproximadamente la mitad de las fijaciones se centran en las áreas de interés más relevantes para la resolución del problema, mientras que en los novatos aproximadamente un tercio de las fijaciones se concentran en esas áreas. En cuanto al rendimiento, se observaron diferencias significativas entre expertos y novatos, mostrando los primeros mejor rendimiento. En cuanto a la duración de la resolución de los ejercicios solo se encontraron diferencias significativas entre expertos y novatos al resolver pseudocódigos. Al tomarse en cuenta como variables de control la inteligencia fluida y la memoria de trabajo de las personas participantes, se descarta que las diferencias observadas se deban a diferencias en las capacidades cognitivas de los participantes.

El que se lleve a cabo un control de las habilidades cognitivas es de relevancia para futuras investigaciones con seguimiento ocular, ya que el contar con métodos adicionales para asegurar la confiabilidad de los datos es importante.

Al disponerse de una batería de ítems más extensa que en investigaciones anteriores, se asegura una mayor confiabilidad externa de los datos, ya que se simula con mayor fiabilidad el tipo de experiencia a la que se verán sometidos los estudiantes en el entorno académico.

Los datos parecen indicar que el uso de diagramas de flujo únicamente es adecuado para ejercicios de baja dificultad, por lo que su uso debería limitarse únicamente a las primeras clases de computación. Para comprender algunos conceptos básicos, como los ciclos.

Los expertos presentaron mayores tiempos de ejecución y menor rendimiento al emplear diagramas de flujo, lo que puede ser indicativo de que se han perdido experticia en el uso de dicho sistema de representación.

El mayor tiempo de ejecución y menor rendimiento en los ejercicios de búsqueda de errores podría ser indicativo de que este tipo de problemas podría demandar una mayor carga de memoria de trabajo, sin embargo, dicha hipótesis no se puso a prueba en esta investigación.

Recomendaciones para docentes de programación

1. En novatos presentar ejemplos en que las áreas más importantes estén representadas de otro color. Pudiéndose utilizar este sistema tanto en los pseudocódigos como en los diagramas de flujo.
2. En novatos variar el sistema de representación de los diagramas de flujo, en función de si forman parte del proceso de inicialización o proceso.
3. A nivel intermedio procurar adoptar lo más rápido posible el uso de pseudocódigos. Dado su mayor similitud a los lenguajes de programación.
4. A nivel intermedio procurar llevar a cabo ejercicios de detección de errores, dado que son más exigentes.

Recomendaciones para investigadores de ciencias de la computación

1. Revisar estudios previos similares hechos con lenguajes de programación específicos. Para determinar en ese lenguaje cuales son las áreas de interés relevantes para resolver problemas.
2. Explorar la posibilidad de llevar a cabo estudios de seguimiento ocular, con el objetivo de contribuir a estudios de control de calidad de software.
3. Dichos estudios se suelen basar en gran parte en la percepción del usuario acerca de la correctitud del sistema.

Recomendaciones para investigadores en ciencias cognitivas

1. Profundizar en las diferencias en la carga en la memoria de trabajo en las tareas de identificación de errores y de reconocimiento de la función que implementa un método.
2. Profundizar en las estrategias cognitivas empleadas por los estudiantes al resolver problemas de algoritmos computacionales. Por ejemplo, desde el modelo de Adquisición – Codificación – Recuperación – Apoyo.

Referencias bibliográficas

- Allaire, J; Xie, Y; Mc Pherson, J; Luraschi, J; Ushley, K; Atkins, A; Wickham, H (2020). Dynamic documents for R.
<https://cran.r-project.org/web/packages/rmarkdown/rmarkdown.pdf>
- Andrzejewska, M; Stolinska, A; Blasiak, W; Peczkowski, P; Rosiek, R; Rozek, B; Sajka, M & Weislo, D (2015). Eye Tracking verification of the strategy used to analyse algorithms expressed in a flowchart and pseudocode. *Interactive learning environments*, 24, 8, 1981-1995.
- Aschwanden, C & Crosby, M (2006). Code scanning patterns in program comprehension. *39 th Annual Hawaii International Conference on Systems*.
- Bednarik, R (2012). Expertise dependent visual attention strategies develop over time during debugging with multiple code representations. *International Journal of human – computer studies*, 70, 143-155.
- Bednarik, R; Busjahn, T; Schulte, C (2014). Eye movements in programming education: analyzing the expert´s gaze. *Proceedings of the First International Workshop*.
- Cardie, C (1999). Integrating case-based learning and cognitive biases for machine learning of natural language. *Journal of experimental & theoretical artificial intelligence*, 11 (3), 297-337.
- Cattell, R & Cattell, A, (2015). Manual prueba factor G de Catell. Madrid: Tea Ediciones.
- Dink, J. W., & Ferguson, B. (2015). eyetrackingR: An R Library for Eye-tracking Data Analysis. Retrieved from <http://www.eyetrackingr.com>.
- Duchowski, A (2007). Eye tracking methodology: theory and practice. South Carolina: Springer.
- Elosúa, M. R., García-Madruga, J. A., Vila, J. O., Gómez-Veiga, I., & Gil, L. (2013). Improving reading comprehension: From metacognitive intervention on strategies to the intervention on working memory executive processes. *Universitas Psychologica*, 12(5), 2-35.
- Gegenfurther, A; Lehtinen, E; Saljo, R (2011). Expertise differences in the comprehension of visualizations: a meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review*, 23, 523-552.

- Lin, Y; Wu, C; Hou, T; Lin, Y; Yang, F; Chang, C (2015). Tracking students cognitive processes during program debugging: an eye movement approach. *IEEE Transactions on education*, 59 (3), 175-186.
- Mason, R; Seton, C & Cooper, G (2016). Applying cognitive load theory to the redesign of a conventional database systems course. *Computer science education*, 26, 1, 68-87.
- Paivio, A (1986). *Mental representations: a dual coding approach*. Oxford. Oxford University Press.
- Pinales, F y Velázquez, C (2014). *Algoritmos resueltos con diagramas de flujo y pseudocódigo*. Aguascalientes: Universidad Autónoma de Aguascalientes.
- Romano, J & Schall, A (2014). *Eye tracking in user experience design*. Boston: Elsevier.
- Rueda, S (2013). *El desarrollo de la comprensión lectora en las carreras de informática*. Departamento de Ciencias e Ingeniería de la Computación. Universidad Nacional del Sur.
- Sampedro, M; Blanco, M; Ponte, D; Leiros, L (2004). *Saliencia perceptiva y atención. La atención: un enfoque pluridisciplinar*.
- SR Research (2017). *EyeLink Data Viewer user's manual*. Ontario. SR Research.
- Tierney, N; Cook, D; McBain, M; Fay, C (2020). Data structures, summaries and visualisations for missing data. <https://cran.r-project.org/web/packages/naniar/naniar.pdf>
- Tsapatsoulis, N (2014). *Learning technologies & open and distance learning*. Nicosia: University of Cyprus.
- Ware, C (2008). *Visual thinking for design*. Burlington: Morgan Kaufmann.
- Wechsler (2012). *WAIS-IV. Escala de inteligencia de Wechsler para adultos IV*.
- Wickham, H & Grolemund, G (2017). *R for data science*. O'Reilly.
- Wickham, H; Chang, W; Henry, L; Pedersen, T; Takahashi, K; Wilke, C; Woo, K (2020). Create elegant data visualizations using the grammar of graphics. <https://cran.r-project.org/web/packages/ggplot2/ggplot2.pdf>
- Wickham, H; Francois, R; Henry, L; Muller, K (2020). *A grammar of data manipulation*. <https://cran.r-project.org/web/packages/dplyr/dplyr.pdf>
- Yi-Chin, L., Yu-Ling, L., & Ying-Shao, H. (2014). The Effects of Meta-Cognitive Instruction on Students' Reading Comprehension in Computerized Reading Contexts: A Quantitative Meta-Analysis. *Journal Of Educational Technology & Society*, 17(4), 186-202.

- Zagermann, J; Pfeil, U & Reiterer, H (2016). Measuring cognitive load using eye tracking technology in visual computing. BELIV 2016. October 24, 2016. Baltimore, USA.