



**INFLUENCIA DE LA INTERACCIÓN ENTRE EL NIVEL DE SATURACIÓN D
INFORMACIÓN EN LOS RESULTADOS DE BÚSQUEDA EN LA WEB Y LA
MOTIVACIÓN AL PENSAMIENTO CRÍTICO SOBRE LA OPERACIÓN DE LO
SESGOS COGNITIVOS DURANTE LA EXPLORACIÓN, SELECCIÓN Y ANÁLISIS
CRÍTICO DE FUENTES DE INFORMACIÓN POR PARTE DE ESTUDIANTES
UNIVERSITARIOS**

***INFLUENCE OF THE INTERACTION BETWEEN THE LEVEL OF INFORMATION
SATURATION IN WEB SEARCH RESULTS AND THE MOTIVATION TO CRITICAL
THINKING ON THE OPERATION OF COGNITIVE BIASES DURING THE
EXPLORATION, SELECTION AND CRITICAL ANALYSIS OF INFORMATION
SOURCES BY UNIVERSITY STUDENTS***

Estudiante:

Martha Vidal Sepúlveda

Dirección de la Tesis:

Profesor Tutor: Dr. Luis Cárcamo Ulloa (Universidad Austral de Chile).

Profesora Tutora: Dra. Magaly Ruiz (Universidad de La Frontera).

Profesora Tutora Externa (Co-Tutela): Dr. Evandro Vieira (Universidad Federal de Río de Janeiro, Brasil)

Fecha de la entrega del documento final de la tesis:

23 de diciembre de 2022

*Dedicada a mis amores, mi tribu:
Cristian, Javiera y Gonzalo.
Y a esa pequeña niña que soñaba con enseñar.*

AGRADECIMIENTOS

Para cerrar un ciclo y emprender nuevos desafíos es necesario agradecer. Este simple, pero sublime rito nos posibilita el tránsito a otros escenarios. Por ello, en estas primeras líneas quiero expresar mi gratitud a todas y todos mis estudiantes que a lo largo de estos años he tenido la fortuna de conocer y dialogar en distintas aulas. Sin duda, han enriquecido mi experiencia docente y forjado mi curiosidad investigadora.

Agradezco a mis directores de tesis Magaly Ruiz de la Universidad de la Frontera y Luis Cárcamo Ulloa de la Universidad Austral de Chile por su colaboración para llegar a fin con esta investigación. Especialmente, corresponder a Luis Cárcamo Ulloa, quién desde el primer correo electrónico, sin siquiera conocernos, mostró respeto e interés por mi proyecto de investigación. Agradezco especialmente su cordialidad, su franqueza, su confianza y la colaboración con mi proyecto de investigación y; el permanente respaldo que me ha brindado, incluso, invitándome a formar parte de uno de sus proyectos de investigación.

Asimismo, dar las gracias a Cristian Olivares Rodríguez, académico de la Facultad de Ingeniería de la Universidad Austral de Chile por su apoyo incondicional en la realización de esta investigación. Cristian, es el creador de la plataforma GoNSA2 utilizada en esta investigación y, además, es coautor de varias de mis publicaciones. Por lo que esta tesis marca el inicio de una colaboración conjunta en una línea de investigación que cruza las disciplinas de comunicación, educación y tecnología.

En este espacio, también agradezco a todos los académicos y académicas con los que tuve la oportunidad de interactuar durante el primer y segundo año, tanto de manera presencial como virtual. Sería inconveniente dar nombres, podría olvidar algunos, pero sepan que atesoro esas discusiones. Asimismo, recuerdo aquí a mis compañeros y compañeras de generación. Juntos formamos un grupo humano diverso e interesante y compartimos parte importante de la historia reciente de nuestro país: un estallido social y una pandemia ¡Cómo olvidar esas reflexiones!

Por supuesto, también reconocer a la Universidad Austral de Chile por la exención de arancel del primer año de estudios doctorales y al Comité de la Beca Eleazar Huerta de la Facultad de Filosofía y Humanidades de la Universidad Austral de Chile por su apoyo con el financiamiento de la mitad del arancel de matrícula durante los tres años restantes. Ese apoyo,

me permitió, pese a las dificultades, continuar en el programa. Y a Karin Osses, quién cumple una labor fundamental de orientación y apoyo administrativo para todos los estudiantes de posgrado de nuestra facultad. Asimismo, al Instituto de Informática de la Universidad Austral de Chile por brindarme el espacio para desarrollar mi investigación a través de la creación y ejecución de un curso de formación general para los estudiantes de primer ingreso.

Finalmente, agradezco infinitamente a mi familia, a mi tribu, por su amor, apoyo constante y paciencia con esta mujer/madre/estudiante que robó horas al día para intentar, al menos, estar lo menos ausente posible ¡Gracias!

Tabla de contenidos

INTRODUCCIÓN

Introducción general.....	9
PRESENTACIÓN Y CONTEXTUALIZACIÓN DE LA INVESTIGACIÓN	13
1. Problema de investigación.....	13
1.1 El aumento de la indexación y su relación con la saturación de información.....	14
MARCO TEÓRICO Y DE ANTECEDENTES	17
1. Educomunicación	17
1.1 La educomunicación en la era de la información	20
2. La gobernanza algorítmica	25
2.2. La búsqueda de información en internet.....	30
3. Pensamiento crítico en la búsqueda de información	32
3.1 Motivación: la clave del pensamiento crítico	34
4. Sesgos cognitivos	35
4.1 Sesgos cognitivos en la búsqueda de información	36
OBJETIVO DE LA INVESTIGACIÓN	37
1. Objetivo general	40
2. Objetivos.....	40
DISEÑO METODOLÓGICO	40
1. Primer estudio cuasi- experimental	41
1.1 Muestra.....	41
1.2 Estrategia de recolección de datos	41
1.3 Análisis de datos	42
1.4. Tareas de búsqueda.....	42
1.5 Protocolo de investigación	43
2. Segundo estudio experimental.....	44
2.1. Variables.....	44
2.2 Muestra	45
2.3 Estrategia de recolección de datos.....	45
2.4 Análisis de datos.....	47
2.5 Tareas de búsqueda.....	48
2.6 Protocolo de investigación	52
2.7 Reglas de análisis	52

3. Plataforma GoNSA2.....	54
Capítulo Primero	60
¿QUIÉN ELIGE MI DIETA INFORMATIVA? LA DATIFICACIÓN AUTOMATIZADA DE PERSONAS EN EL AMBIENTE DIGITAL	60
1. Resumen	60
2. Introducción.....	61
3. La datificación de personas como insumo.....	61
4. La conformación de subjetividades en el ambiente digital.....	62
5. De la subjetivación del cuerpo a las subjetividades digitales.....	64
6. Los motores de búsqueda de información como herramienta para la creación de subjetividades	66
7. Conclusión	69
Capítulo Segundo	72
WHAT KIND OF SOURCES DO I NEED?: CRITICAL SEARCH FOR INFORMATION ON THE WEB.....	72
1. Abstract.....	72
2. Introduction	73
3. Algorithmic challenges for autonomous learning	74
3.1 Critical thinking as a learning competence.....	74
3.2 Information search for autonomous learning.....	77
4. Material and methods	79
4.1 Participants	79
4.2 Design of information search tasks	79
4.3 Materials	80
4.4 Procedure	81
4.5 Source categorization	81
5. Results	82
5.1 Quality of the information sources according to their process of content edition..	82
5.2 Website ranking.....	83
6. Discussion and outlook.....	85
Capítulo Tercero	89
BUSCAR Y SELECCIONAR INFORMACIÓN EN INTERNET: SESGOS, SATURACIÓN DE INFORMACIÓN Y PENSAMIENTO CRÍTICO.....	89
1. Resumen	89
2. Introducción.....	90

3. Material y métodos	94
4. Resultados.....	99
4.1. Sesgo de anclaje	99
4.2 Sesgo de ranking.....	109
4.3 Sesgo de disponibilidad de resultados	111
5. Conclusiones.....	113
6. Limitaciones y proyecciones	115
Capítulo Cuarto	117
“BÁSICAMENTE, TODO LO PUEDES ENCONTRAR AHÍ: CREENCIAS, SESGOS Y ESTRATEGIAS DE LOS ESTUDIANTES UNIVERSITARIOS EN LOS MOTORES DE BÚSQUEDA”	117
1. Resumen	117
2. Introducción.....	118
4. Metodología.....	121
4.1 Participantes	121
4.2 Muestra	121
4.3 Procedimiento	122
5. Resultados.....	124
6. Discusión	147
7. Conclusiones.....	151
8. Limitaciones	152
CONCLUSIONES.....	153
1. Conclusiones.....	153
2. Implicancias.....	158
3. Líneas de trabajo futuras	159
REFERENCIAS	161

Índice de tablas

TABLA 1: TAREAS DE BÚSQUEDA EN INTERNET CUASIEXPERIMENTO	43
TABLA 2: PROPUESTA DE TAREAS DE BÚSQUEDA EN INTERNET	49
TABLA 3: CUESTIONARIO DE EVALUACIÓN DEL JUICIO DE EXPERTOS	50
TABLA 4: REGLAS DE ANÁLISIS	53
TABLA 5: INFORMATION SEARCH TASKS	80
TABLA 6: TAREAS DE BÚSQUEDA DE INFORMACIÓN	97
TABLA 7: SESGO DE ANCLAJE ENTRE LA CONSIGNA - CONSULTAS	101
TABLA 8 : SESGO DE ANCLAJE ENTRE LA CONSULTA- TÍTULO DE LA BIBLIOTECA	103
TABLA 9: SESGO DE ANCLAJE ENTRE LA CONSULTA - SNIPPET DE LA BIBLIOTECA	104
TABLA 10: SESGO DE ANCLAJE ENTRE EL RESUMEN - SNIPPETS DE LA BIBLIOTECA	106
TABLA 11: SESGO DE ANCLAJE ENTRE CONSULTA - TÍTULO	107
TABLA 12: SESGO DE ANCLAJE ENTRE LA CONSULTA - SNIPPET	108
TABLA 13: TAREAS DE BÚSQUEDA	122
TABLA 14: CATEGORÍAS PREESTABLECIDAS	123
TABLA 15: CATEGORÍAS EMERGENTES	124
TABLA 16: TRAYECTORIA TAREA "COVID 19"	144
TABLA 17: TRAYECTORIA TAREA "ELECCIONES 2021"	145
TABLA 18: TRAYECTORIA TAREA "NOTICIAS FALSAS"	146

Índice de ilustraciones

ILUSTRACIÓN 1: RESULTADOS EVALUACIÓN EXPERTOS	51
ILUSTRACIÓN 2: DISEÑO CARACTERIZACIÓN DE SESGOS COGNITIVOS EN LA BÚSQUEDA DE INFORMACIÓN.....	54
ILUSTRACIÓN 3: INTERFAZ DE INICIO GoNSA2	55
ILUSTRACIÓN 4: INTERFAZ DE INICIO DE SESIÓN GoNSA2.....	55
ILUSTRACIÓN 5: INTERFAZ DE INICIO	56
ILUSTRACIÓN 6: INTERFAZ TAREAS GoNSA2	56
ILUSTRACIÓN 7: INTERFAZ TAREA COVID-19	57
ILUSTRACIÓN 8: INTERFAZ DE BÚSQUEDA GoNSA2	58
ILUSTRACIÓN 9: INTERFAZ GUARDAR RESPUESTA GoNSA2	58
ILUSTRACIÓN 10: INTERFAZ RESPUESTA ENVIADA GoNSA2	59
ILUSTRACIÓN 11: GoNSA2, TASK RESOLUTION INTERFACE.....	81
ILUSTRACIÓN 12: SOURCE TYPE CHARACTERIZATION	83
ILUSTRACIÓN 13: VARIATION OF THE SERPs FOR THE TASK: ‘HIV IN CHILE’	84
ILUSTRACIÓN 14: VARIATION OF THE SERPs FOR THE TASK: ‘HOW TO BUILD A CAR’	84
ILUSTRACIÓN 15: SERP VARIATION IN THE 13 QUERIES PERFORMED BY A STUDENT TO SOLVE A RESEARCH / EXPLORATORY TASK	85
ILUSTRACIÓN 16: GoNSA2: INTERFAZ DE RESOLUCIÓN DE TAREAS DE INFORMACIÓN	98
ILUSTRACIÓN 17: DISEÑO DE LA METODOLOGÍA DE CARACTERIZACIÓN DE SESGOS COGNITIVOS.....	99
ILUSTRACIÓN 18: DISTRIBUCIÓN DE LA POSICIÓN DE LA SNIPPET DE LA BIBLIOTECA PERSONAL	110
ILUSTRACIÓN 19: CANTIDAD DE CONSULTAS POR TAREA	112
ILUSTRACIÓN 20: MAPA DE ENTREVISTAS SEMIESTRUCTURADAS	131
ILUSTRACIÓN 21: RESULTADOS ESCALA EMPC	142

Introducción general

Los primeros cuatro capítulos de esta investigación se componen de una discusión teórica y dos acercamientos empíricos al tema de estudio. El capítulo cinco corresponde a la presentación de las conclusiones generales del estudio.

En el capítulo 1 titulado “¿Quién elige mi dieta informativa? La datificación de personas en el ambiente digital” se plantea la discusión teórica de las dificultades que enfrentamos en el habitar en el ambiente digital. Primero como espacio de flujo de información y, segundo, como creador de identidades digitales. En este contexto se discute el lugar de los motores de búsqueda como mecanismo automático colaborador en la creación de subjetividades. Se reflexiona respecto de las implicancias de las decisiones de los algoritmos de los motores de búsqueda. Específicamente, sobre la construcción de perfiles de usuarios que conlleva a la conformación de identidades digitales que nos clasifican y agrupan para determinar nuestros posibles intereses. Este capítulo sienta las bases de las observaciones a considerar en los estudios empíricos con estudiantes universitarios. Esta reflexión teórica problematiza respecto de las competencias críticas de las personas para sobreponerse a la influencia de la personalización algorítmica, vale entonces preguntarse si ¿Comprendemos las implicancias de dicha personalización en la conformación de nuestra dieta informativa?

¿En el capítulo 2 “What kind of sources do I need?: critical search for information on the web”, buscamos indicios respecto de las fuentes de información accesibles mediante una búsqueda en la web. En los últimos años la web crece a un ritmo inabarcable para un ser humano, por lo que buscamos determinar empíricamente la calidad de edición que los usuarios hacen de las fuentes de información indexadas en la nube. En este capítulo se detallan los avances en la literatura para describir los numerosos obstáculos que enfrentan, muchas veces sin saberlo, los estudiantes en el momento de realizar una búsqueda de información. El objetivo es determinar la calidad de las fuentes de información de las SERP (SERP, Search, Engine Result Page) que reciben los estudiantes a partir de sus consultas, considerando los sesgos aplicados por los algoritmos de los sistemas de búsqueda. En ese contexto nos preguntamos ¿Los estudiantes tienen la oportunidad de acceder a información

de calidad a partir de sus consultas en la web? ¿Las competencias críticas de los estudiantes facilitan el acceso a una mayor diversidad en sus consultas en la web? Se eligió el diseño de un cuasi- experimento en el cual los estudiantes realizaron búsquedas de información en la plataforma GoNSA2. Para las tareas se utilizó un diseño y tópicos previamente validados en la literatura. La muestra se compuso de 53 estudiantes de primer año de ingeniería civil informática, de los cuales 5 corresponden al género femenino, lo que refleja la distribución por género de la matrícula en dicha carrera. Como resultado principal se evidencia que la elaboración del listado de resultados no considera el tipo de edición de las fuentes, sino que dependen de otras variables descritas en estudios anteriores que se detallan en el capítulo. Sin embargo, la mayor parte de las fuentes alcanzadas por los estudiantes corresponde a información que declara un proceso de edición, pero con alta presencia de información no validada. Se concluye que el aumento de la información con proceso editorial estaría asociado con la presencia del tópico de la tarea en la agenda mediática. No obstante, cabe mencionar que este estudio se limita a determinar que la fuente de información declare un proceso de edición y, por tanto, no se analiza la veracidad del contenido de la página web.

En el capítulo 3 “Buscar y seleccionar información en internet: Sesgos, saturación de información y pensamiento crítico, se profundiza sobre los sesgos inherentes en internet”, tal como revisamos en los capítulos 1 y 2 los filtros algorítmicos trajeron consigo ventajas para acceder a la información en menor tiempo, pero a su vez implantaron sesgos. En este apartado profundizamos en el análisis de la literatura respecto del origen y consecuencias de los sesgos presentes en internet. Hasta nuestro conocimiento hemos encontrado estudios que analizan los sesgos en internet mediante técnicas de *scraping*, modelado de usuarios y experimentos que manipulan el listado de resultados. En concreto, los sesgos presentes en los resultados de búsqueda han sido descritos en la literatura como una de las principales barreras en una búsqueda de información. Por ejemplo, se ha establecido que las personas son influenciadas por el orden de los resultados propuesto por el buscador.

Por lo anterior, en esta investigación se analizan las búsquedas como un proceso individual mediado por la tecnología, en el cual la persona explora, selecciona y analiza la información, poniendo sobre cada decisión sus representaciones culturales, involucrando sus emociones y acciones. Nuestro objeto de estudio corresponde a la operación de los sesgos

cognitivos sobre la exploración y selección de información en la web para la resolución de tareas de información en estudiantes universitarios. En esta investigación buscamos responder ¿Cuáles son los sesgos que operan durante la búsqueda de información para resolver una tarea académica? ¿La motivación de los estudiantes a pensar críticamente influye significativamente sobre la operación de sesgos cognitivos? ¿Los estudiantes que reciben una lista de resultados saturada manifiestan un comportamiento más sesgado?

En esta línea de trabajo buscamos evidencias de la operación de sesgos cognitivos en las búsquedas de información en internet, considerando la saturación de información en la web y la motivación al pensamiento crítico como variables del estudio. Con esa finalidad diseñamos un experimento de carácter mixto cuantitativo-cualitativo secuencial. En el cual se buscó ampliar la muestra por género y área de estudio, incluyendo a las carreras de ingeniería civil informática y periodismo. Participaron 64 estudiantes de primer año, 42 de género masculino y 22 de género femenino. En este experimento se elaboraron tareas *ad hoc*, las que fueron evaluadas por un comité experto. De acuerdo con los resultados del primer cuasi experimento se incluyeron tareas con tópicos presentes en la agenda mediática nacional como medida de control de esta variable. Nuevamente utilizamos la plataforma GoNSA2 para recopilar las trazas de comportamiento de búsqueda de los participantes. Para analizar las trazas de información se creó una estrategia de diseño de análisis propia que se detalla en el capítulo 3. Este diseño nos permitió observar y caracterizar la operación de los sesgos de anclaje, ranking y disponibilidad de resultados en distintas instancias del proceso de búsqueda. Además, se determinó que la saturación de información en interacción con la motivación al pensamiento crítico influye significativamente en la ocurrencia de sesgos.

El capítulo 4 profundiza el análisis de este segundo acercamiento empírico. Este capítulo titulado “Básicamente, todo lo puedes encontrar ahí: Percepciones, sesgos y estrategias de los estudiantes universitarios en los motores de búsqueda” busca indagar respecto de la coherencia entre el comportamiento de los participantes durante la búsqueda de información en internet y las respuestas en sus entrevistas semiestructuradas. Para ello, se realizaron entrevistas semiestructuradas a 13 estudiantes universitarios que previamente participaron de la sesión experimental. Con este fin, se contrastaron los datos de los reportes del comportamiento de búsqueda con las respuestas de las entrevistas para identificar los

usos, percepciones, criterios de selección de páginas webs y la operación de sesgos cognitivos en la búsqueda de información. El principal hallazgo es la observación del sesgo de familiaridad con Google, es decir, la tendencia a confiar en la selección de dicho motor de búsqueda. Asimismo, también se constata la operación de los sesgos cognitivos de información y de confirmación.

Finalmente, el capítulo 5 corresponde a las conclusiones generales del estudio, implicancias y líneas futuras de investigación.

PRESENTACIÓN Y CONTEXTUALIZACIÓN DE LA INVESTIGACIÓN

1. Problema de investigación

El motor de búsqueda es un dispositivo tecnológico de uso masivo; sin embargo, su origen data de poco más de 30 años. La creación de la World Wide Web posibilitó la oportunidad de avanzar en el mejoramiento de los algoritmos de indexación de páginas web, dando paso al nacimiento de los buscadores. Los motores de búsqueda crecieron juntamente con la masificación y la ampliación de la World Wide Web. Esta fue posible por los avances de la web 2.0 que facilitaron la creación de contenido y el aumento de la indexación de páginas web. Altavista, es reconocido como el primer motor de búsqueda que permitió la emisión de consultas en lenguaje natural, incluyendo varias lenguas, y el que más se acerca a los buscadores que conocemos hoy.

Los motores de búsqueda se pueden clasificar en generalistas y especializados. Los generalistas reúnen todo tipo de información, sin un criterio de validación conocido respecto de la edición de los contenidos. Los más populares son Google, Bing, Aol, Duckduckgo, entre muchos otros. En cambio, los buscadores especializados basan su estrategia de indexación en la aplicación de criterios rigurosos para evaluar las fuentes de información. De los buscadores académicos destacan: Google Scholar, Live Search Academic y Scirus.

Estos sistemas de búsqueda gozan de popularidad en la educación por la facilidad de acceder a gran cantidad de información a partir de la emisión de una consulta en lenguaje natural, presentan un alto nivel de satisfacción por parte de los usuarios y se encuentran integrados en los principales navegadores web de la actualidad. Estos sistemas de recuperación de información permiten acceder, de manera rápida y sencilla, a una gran base de conocimiento social de gran utilidad para la investigación y la educación. Sin embargo, es importante establecer los límites en la calidad de la información a la cual es posible acceder por medio de estos dispositivos tecnológicos.

Si bien el acceso a la web no es un derecho, se ha observado un aumento de las conexiones a internet en los hogares y los centros educativos que han posibilitado la ubicuidad del acceso a la información, a través de distintos dispositivos a todo tipo de

documentos web listos para ser descargados, imponiendo nuevos desafíos educativos entre quienes que cuentan con este privilegio y desafíos sociales entre los gobiernos para avanzar en la reducción de la brecha de acceso.

1.1 El aumento de la indexación y su relación con la saturación de información

Actualmente, es posible que estados, organizaciones de todo tipo, incluso, personas naturales creen y compartan contenido. Esto debido a la facilidad, el bajo costo, la rápida distribución y visibilización en diferentes interfaces que ha propiciado el aumento vertiginoso del contenido en la web. En la literatura se describe una mediamorfosis ahogada por desórdenes informativos aparentemente estructurales, tales como la sobresaturación informativa y la infoxicación (Aguaded y Romero-Rodríguez, 2015; Romero-Rodríguez, L. M. y Mancinas R., 2016; Romero-Rodríguez y Aguaded, 2016). Otros como Cobo (2011) se manifiestan desesperanzados y la denominan una explosión de desinformación, que confunde y que es imposible de digerir por las personas.

La saturación de información se hizo realidad en los inicios de la sociedad de la información, pero alcanzó su crisis con el advenimiento de la web 2.0. En la literatura diferentes autores han conceptualizado la saturación de información como un fenómeno crítico en la web 2.0, con síntomas como agotamiento físico y mental provocados por la sobre exposición y consumo excesivo de información.

En definitiva, convivimos con esta sobrecarga de información generada por la capacidad de la tecnología de emitir cada vez más información en menos tiempo. Sin embargo, la capacidad cognitiva de las personas es limitada y no puede dedicar el tiempo necesario para gestionar adecuadamente toda la información que recibe. Las personas nos vemos afectadas cuando el exceso de datos de entrada a nuestro sistema (la mente) sobrepasa la capacidad de procesamiento. Lo anterior, afecta la calidad en la toma de decisiones (Speier et al., 1999). Asimismo, influye sobre la atención, dificultando su capacidad para procesar las piezas informativas que sobrecargan nuestra memoria operativa (Rivas Navarro, 2008). Según Cornellá (2004) para alcanzar a procesar con rapidez la información es necesario tener un conocimiento previo nivel experto del tema.

Actualmente, la sobrecarga de información pone a prueba nuestra capacidad tanto de aprender de manera autónoma como de tomar decisiones, lo que da paso a una paradoja de la información. Somos una sociedad rica en información y, asumimos que gran parte de la información a la que accedemos es de calidad. Sin embargo, se ha demostrado que existe una creciente cantidad de datos tóxicos (“data smog”) que aumentan el tiempo que utilizamos analizando y/o procrastinando frente a pilas de información que nos seducen desde distintos medios (Shenk, 2003). Este ecosistema informativo ha sido descrito como “economía de la atención” (Goldhaber, 1997; Davenport & Beck, 2001), porque existen demasiadas fuentes de información que compiten por la atención de las personas.

De este fenómeno también son parte, los estudiantes universitarios que participan activamente del ambiente digital como consumidores y productores de altas dosis de información. Según Serrano-Puche (2014), en este flujo constante de información digital la lectura se vuelve más apresurada, el pensamiento más distraído, disminuyendo la concentración en la lectura y el análisis profundo de la información, “nos hemos habituado al surféo informativo” (Serrano-Puche, 2013, p. 35). Esto refuerza las complejidades ambientales del aprendizaje autónomo y profundo en la sociedad por la excesiva información.

Por otra parte, el incremento de las habilidades de uso operativo de las aplicaciones tecnológicas por parte de las personas les permite acceder a grandes volúmenes de información. Este fenómeno ha generado una serie de comportamientos entre los usuarios que buscan afrontar la sobrecarga de información a la que nos vemos enfrentados a diario. Por ejemplo, se ha visto una tendencia por la selección de fuentes de información que refuerza sus opiniones, como una medida de abordar la infosaturación. Este fenómeno parece estar potenciado por las decisiones algorítmicas detrás de las tecnologías disruptivas, que seleccionan y filtran la información para los usuarios a partir de su perfilado digital, desconocido por muchos usuarios, pero percibido como un soporte tecnológico al abordaje de la sobrecarga de información (Dias, 2014).

En consecuencia, cada vez que un estudiante inicia una búsqueda en internet se disparan una serie de procesos algorítmicos y cognitivos que interactúan, se potencian y cruzan el espacio digital para disponer de documentos creados por agentes humanos o algorítmicos (*bots*). En esta interacción humano-computador, los estudiantes y los algoritmos

toman decisiones en función de la gran cantidad de información que es necesario procesar. Sin embargo, los agentes humanos reciben más información de la que pueden gestionar y tienden a confiar en la clasificación propuesta por el algoritmo del sistema de búsqueda (Noble, 2018), los que actúan de manera implícita y opaca (Pasquale, 2015; Kitchin, 2017). Cabe preguntarse si las personas ¿Somos conscientes de la selección algorítmica de información? ¿Entendemos el alcance de la selección de nuestra dieta informativa?

MARCO TEÓRICO Y DE ANTECEDENTES

El marco teórico de esta tesis se inserta en los estudios de la educomunicación (Kaplún, 1998), la gobernanza algorítmica (Pasquale, 2015; Kitchin, 2017; Danaher et al., 2017), la teoría de sesgos cognitivos (Tversky y Kahneman, 1983); y su relación con el pensamiento crítico.

1. Educomunicación

La educomunicación es un campo de estudios de la comunicación que reúne a la educación y la comunicación. Como campo aborda las dimensiones teórico-prácticas de ambas disciplinas. Como detallaremos en este apartado, no existe acuerdo respecto del origen de la educomunicación; sin embargo, se han identificado a distintos referentes como Célestin Freinet, Frank Raymonds y Denis Thompson, en el ámbito anglosajón, y a Mario Kaplún y Paulo Freire, en Latinoamérica. De acuerdo con Kaplún (1998), Freinet fue el primer educador que utilizó los medios de comunicación como material didáctico en las aulas escolares en la década del 20 del pasado siglo. Su principal aporte fue el cambio de paradigma educativo, desde un aprendizaje conductista, basado en el aprendizaje memorístico, hacia un paradigma socio constructivista en el cual los educandos participan de la construcción de sus aprendizajes mediante la práctica reflexiva/dialógica. En concreto, los estudiantes aprendían mientras elaboraban un periódico escolar. Esta metodología del periódico escolar alcanzó gran popularidad y se extendió hacia distintas escuelas francesas y españolas. Pero, según Masterman (1933) la historia de la educomunicación parte en Gran Bretaña con la publicación de “Culture and environment” de Leavis y Thompson. Este texto habría sentado las bases de lo que más tarde conoceríamos como alfabetización mediática. En dicho ejemplar se alerta sobre la influencia negativa de los mass media en la sociedad inglesa. Los acusaban de ser parte importante de la decadencia social y moral, por lo que emplazaban, específicamente, a los profesores de lengua a modificar su práctica educativa para preservar los valores tradicionales.

De acuerdo con Aparici (2011) su origen es Latinoamericano y sus principios surgen de la pedagogía crítica de Paulo Freire y desde allí llegaría a España. Para este autor la reunión de expertos en Santiago de Chile convocados por UNICEF, la UNESCO y CENECA

(Centro de Indagación de Indagación e innovación artística) es el hito que marca una de las principales definiciones del campo:

“educación incluye, sin reducirse, el conocimiento de los múltiples lenguajes y medios por los que realiza la comunicación personal, grupal y social. Abarca también la formación de sentido crítico, inteligente, frente a los procesos comunicativos y sus mensajes para descubrir los valores culturales propios y la verdad.” (CENECA, UNICEF, UNESCO, 1992, citado en Aparici, 2011)

Ahora bien, más allá de la discusión del origen de la educación, sin duda, el desarrollo del campo de estudio requirió de la convergencia de otras perspectivas, tales como: la Escuela de Frankfurt con sus estudios críticos de las industrias culturales, las teorías comunicativas de Marshall McLuhan, la comunicación para el desarrollo de América Latina y los aportes de reconocidos educadores como Paulo Freire y Mario Kaplún, entre otros. En este contexto, a partir de la década del 70, los estudios sobre educación se extendieron a distintos territorios, pero con diferentes perspectivas que definieron el campo como “(...) un territorio heterogéneo y plural. (Coslado, 2012)”.

Luego de establecido el campo, comienza el camino de la institucionalización de la educación. En 1973 el Consejo Internacional de Cine y Televisión (CICT) elabora una de sus primeras definiciones:

“Por educación en materia de comunicación cabe entender el estudio, la enseñanza y el aprendizaje de los medios modernos de comunicación y de expresión a los que se considera parte integrante de una esfera de conocimientos específica y autónoma en la teoría y en la práctica pedagógica, a diferencia de su utilización como medios auxiliares para la enseñanza y el aprendizaje en otras esferas del conocimiento como las matemáticas, la ciencia y la geografía.” (Morsy, 1984, p.7)

La institucionalización de la educación se alcanza por el interés manifestado por la UNESCO en la década del 70' respecto de la elaboración de políticas de comunicación que promovieran el acceso de los pueblos, sobre todo los más empobrecidos, a los mensajes de los mass media y, con ello, aumentarían y mejorarían el uso de dichos medios para la creación de programas educativos y culturales. En esta misma línea, la UNESCO en sucesivos encuentros elabora informes de expertos que buscan establecer un nuevo orden mundial “Hacia un nuevo orden mundial de información y comunicación, más justo y más eficiente”, tal como explicitaba el informe McBride (1980).

La consolidación del campo de estudios se alcanza en 1984 con la edición del texto *La educación en materia de educación*, que reunía a diferentes autores que establecen las líneas y retos de los estudios de educomunicación, definiéndolo como un campo teórico-práctico. Según este texto las principales tareas son: incrementar el uso de los medios de comunicación en la educación, el estudio del mensaje de los medios y la enseñanza de las técnicas y tecnologías de los medios (Morsy, 1984).

De esta forma, en la segunda mitad del siglo XX surgen distintos movimientos educomunicativos primero locales, luego regionales y finalmente internacionales. Estos grupos buscaron dar respuesta a sus necesidades culturales y sociales (Aparici, 2011). Este nuevo campo es definido por diversos autores y organizaciones, destaca entre ellas la definición de Mario Kaplún, en su libro *Una pedagogía de la comunicación*, donde señala que la educomunicación:

“(…) tendrá por objetivo fundamental el de *potenciar a los educandos como emisores*, ofreciéndoles posibilidades, estímulos y capacitación para la autogeneración de mensajes. Su principal función será, entonces la de proveer a los grupos educandos *canales y flujos de comunicación* -redes de interlocutores, próximos y distantes- para el intercambio de tales mensajes. Al mismo tiempo, continuará cumpliendo la función de proveedora de materiales de apoyo; pero concebidos ya no como meros transmisores-informadores sino como generadores de diálogo, destinados a activar el análisis, la discusión y la participación de los educandos y no a sustituirlas.” (Kaplún, 1998, p.244)

Esta definición de educomunicación se inserta en una concepción comunicativa-educativa, inserta en la tradición dialógica crítica, con la influencia de los aportes de Paulo Freire en la década del 70’ (Coslado, 2012).

De la tradición educomunicativa es posible diferenciar dos enfoques: el instrumental y el dialógico. El enfoque instrumental se concentra en el mundo anglosajón y hace referencia al manejo de la tecnología. En cambio, el enfoque iberoamericano es un planteamiento dialógico con claras referencias a las pedagogías críticas. De este modo, en el enfoque instrumental la educomunicación es denominada comúnmente como *media literacy* o *media education*. Desde esta perspectiva, el docente asume el rol de experto en tecnologías y es el encargado de alfabetizar a sus estudiantes en el uso de nuevas tecnologías. En cambio, el enfoque dialógico promueve la reflexión crítica del estudiante, fomentando el diálogo y la actitud emancipadora. Este carácter dialógico y participativo requiere no solo de la tecnología

sino también del cambio de actitudes y de concepciones (Aparici, 2011). La educomunicación:

“(…) tendría por finalidad la construcción y creación colectiva a través del intercambio simbólico y el flujo de significados. Esto supone considerar, en primer lugar, la naturaleza colaborativa y participativa de la Educomunicación; en segundo lugar, sus posibilidades creativas y transformadoras y, en tercer lugar, los medios y códigos a través de los que se establece el proceso educomunicativo.” (Coslado, 2012, p.165)

Sin embargo, a pesar de los numerosos trabajos y experiencias educomunicativas críticas, el enfoque predominante en las instituciones educativas, a nivel global, ha sido el enfoque instrumental, promovido principalmente por los gobiernos de turno, los cuales han implantado un modelo educomunicativo afín con la visión neoliberal. Lamentablemente, esta decisión política ha impactado en la formación de los países más empobrecidos. Específicamente, en nuestro país la dictadura impuso el modelo neoliberal en todos los ámbitos de la sociedad, incluida la educación. Esta decisión aumentó la desigualdad educativa y con ello el acceso a una formación educomunicativa dialógica-crítica. Por lo anterior, a pesar del desarrollo de una amplia tradición investigadora crítica, las pedagogías críticas aún no ingresan a las aulas chilenas. Es más, el currículum nacional, incluso con la llegada de la democracia en la década del 90’, preserva el sello neoliberal, limitando a la educomunicación a un enfoque instrumental que se ocupa del número de computadores y de conexiones a internet. Este desolador estado de la formación en competencias mediáticas y digitales, tanto en profesores como en estudiantes, quedó expuesto en Chile y el mundo con la pandemia COVID-19.

1.1 La educomunicación en la era de la información

La UNESCO desde su conferencia de Grünwald en 1982, ha planteado su interés en la educación mediática como una estrategia pedagógica que conduce a la alfabetización mediática. En sus inicios las demandas de alfabetización mediática se concentraban de manera casi exclusiva, en el contexto de educación formal. Sin embargo, con los años esta exigencia se ha ido deslocalizando, posiblemente, por los retos que impone la llegada de la web 2.0, lo que no ha variado es su constante promoción. Así, en su informe del año 2005, la UNESCO establece un nuevo concepto de “alfabetización informativa” plantea el concepto

de sociedades del conocimiento para incluir las dimensiones sociales, éticas y políticas. Este concepto engloba a todas las formas posibles de conocimiento y cultura que participan en la formación de sociedades, reconociendo en ellas la influencia del progreso científico y la tecnología. Este nuevo concepto pretende contrarrestar el avance de la conformación de una única sociedad posible basada en el progreso tecnológico. De acuerdo con su visión, el desarrollo tecnológico debe propiciar el desarrollo de sociedades del conocimiento que tengan como finalidad avanzar hacia el desarrollo humano basado en sus derechos (Varis y Pérez Tornero, 2012). En dicho informe se plantea que la complejidad propia de la sociedad de la información y su cultura mediática exigen formar una conciencia mediática y humana, para lo cual es necesario refundar el sistema de medios a partir de la promoción de la alfabetización mediática. Pero, esta alfabetización mediática debe ser útil para formar una conciencia mediática ciudadana, para ello debe considerar:

- “Acceso y apropiación de la tecnología (herramientas, códigos y lenguajes) para ser capaces de recibir, crear y difundir información, facilitando su participación activa en la sociedad.
- Perspectiva crítica que facilite la evaluación y selección de información y que, además, lo habilite para cuestionar el desarrollo tecnológico.
- Fomento del intercambio y la comprensión mutua entre culturas, respecto de la consecución de derechos y responsabilidades universales.” (Varis y Pérez Tornero, 2012, p. 83).

En este sentido, la UNESCO (2005) promueve el desarrollo de la alfabetización informativa como:

“capacidad para identificar, producir, tratar, transformar, difundir y utilizar la información con vistas a crear y aplicar los conocimientos necesarios para el desarrollo humano. Estas sociedades se basan en una visión de la sociedad propicia la autonomía y engloban las nociones de pluralidad, integración, solidaridad y participación.” (Varis y Pérez Tornero, 2012, p.95)

Esta conceptualización plantea que es necesario proveer a las personas de las competencias críticas y herramientas necesarias para desenvolverse como ciudadanos autónomos capaces de realizar un consumo crítico de los medios y la información disponible (Buckingham, 2007; 2008; Wilson et al., 2011).

A la fecha, la alfabetización mediática se ha abordado desde dos enfoques opuestos: promover y proteger (von Feilitzen, 2003). En primer lugar, desde la promoción de las habilidades asociadas al análisis y reflexión de los mensajes enviados desde los medios, para lo cual es necesario promover esta habilidad para formar ciudadanos activos en la construcción de una democracia participativa. No obstante, para su éxito se propone que su formación debiese iniciarse desde la primera infancia como parte de la educación formal, pero aún existen pocas iniciativas que vayan en esa dirección (Buckingham y Domaille, 2003). En segundo lugar, desde la protección de los ciudadanos más vulnerables a la sobrecarga de información que imponen los medios y que gatillan tanto nuestro modo automático como nuestra conexión indirecta con los medios (Potter, 2013). Sin embargo, se ha observado que el proteccionismo mediático, en caso de existir compromiso de los gobiernos, ha carecido de implicancias prácticas debido a la falta de fiscalización (Barilá, 2003). Buckingham (2008) coincide en que el desarrollo de la alfabetización mediática involucra el desarrollo del pensamiento crítico y con ello la formación de ciudadanos más activos, por tanto, no ha contado con el interés ni el financiamiento suficiente de los gobiernos, especialmente en países más empobrecidos.

Actualmente, se proponen modelos intermedios de cooperación entre los principales actores de la alfabetización mediática; los ciudadanos y las familias como responsables primarios de la formación crítica frente a la invasión de los medios; los gobiernos en la generación y fiscalización de políticas públicas que controlen los mensajes emitidos a través de los distintos medios que operan en la actualidad (Clark, 2003); y la industria mediática como generadora de mensajes que buscan atraer audiencias como su principal propósito, pero que se comprometen con espacios para la formación de audiencias críticas (Shipard, 2003).

En ese sentido Ferrés y Piscitelli (2012) articula una completa propuesta de alfabetización mediática en la cual incorpora conocimientos, destrezas y actitudes relacionados con las dimensiones de: lenguajes, tecnología, procesos de interacción, procesos de producción y difusión, ideología, valores y, estética. En cada uno de ellos establece indicadores en el ámbito de análisis; respecto a las personas que reciben e interaccionan con mensajes y, en el ámbito de la expresión; como personas que producen mensajes.

En esta línea en el año 2011, la UNESCO presenta el currículum en alfabetización mediática e informacional (AMI) para profesores (Wilson et al., 2011), siendo el marco de referencia internacional para el desarrollo de competencias para educadores. En este informe se declara a la alfabetización mediática pone en el centro habilidad para entender las funciones de los medios, para evaluar el desempeño de sus funciones y el compromiso con los medios para la autoexpresión. Asimismo, entrega los elementos clave para la alfabetización mediática:

1. Entender el papel y las funciones de los medios en las sociedades democráticas.
2. Entender las condiciones bajo las cuales los medios pueden cumplir sus funciones.
3. Evaluar de una manera crítica el contenido de los medios a la luz de las funciones de los medios.
4. Comprometerse con los medios para la autoexpresión y participación democrática.
5. Revisar destrezas (incluyendo TICs) para producir contenido.

No obstante, los avances de las nuevas generaciones en el uso y apropiación de los nuevos medios y plataformas han obligado a repensar el concepto de alfabetización mediática tradicional para incorporar estos nuevos escenarios de aprendizaje y colaboración. Siemens (2006) propone el conectivismo como un nuevo paradigma según el cual las personas aprendemos a partir de diferentes conexiones de nodos y del intercambio de información entre los nodos. Estos nodos pueden ser personas, contenidos, organizaciones, entre otros. Al conectivismo se le ha relacionado con el modelo rizomático propuesto por Deleuze y Guattari (1987), donde no existe un centro, sino que cualquier parte de la estructura puede generar nuevos elementos. Este modelo de rizoma se ha aplicado al aprendizaje -entendiendo que la conformación del currículum no depende exclusivamente de los expertos- como un proceso de negociación en el tiempo con todos los participantes del proceso de aprendizaje (Cormier, 2008)

Cope y Kalantzis (2010) destacan que la ubicuidad de las tecnologías y las redes sociales han facilitado la conformación de un nuevo paradigma educativo, basado en la posibilidad de aprender en cualquier momento y lugar. Este impacto en la vida cotidiana

implica que los educandos integren el aprendizaje en su rutina diaria (Burbules, 2012). Ahora bien, esta nueva forma de aprender ha modificado las rutinas, las fuentes de información y los espacios de aprendizaje, impactando en los hábitos de lectura y de búsqueda de información, entre otras áreas. También, surgen nuevos retos, como el aprendizaje a lo largo de la vida, tanto en contextos formales e informales. En definitiva, estos nuevos escenarios de aprendizaje pueden ser intencionales o casuales, pero ponen el acento en la necesidad del autoaprendizaje, donde la persona controla su proceso de manera autónoma (Alexander et. al., 2004), detectando sus necesidades y estableciendo sus propios objetivos de aprendizaje.

Esta deslocalización de los aprendizajes implica que los estudiantes “aprenden cosas” de manera independiente que bien podrían complementar la educación formal. Sin embargo, aún existe un vacío que impide visibilizar los aprendizajes no formales e informales a los espacios formales. Cobo y Moravec (2011) lo denominan “aprendizaje invisible”, su propuesta busca integrar al currículum educativo prácticas pertinentes que integren estos aprendizajes para reducir las brechas entre la enseñanza formal y las necesidades del mundo laboral, tales como el autoaprendizaje y la adaptabilidad.

En este mismo sentido, se han propuesto nuevas literacidades basadas en la nueva ecología de medios que se alejan de la idea de proteger a los niños y jóvenes de los efectos de los medios. Autores como Scolari (2016) plantean que no es posible mantener el concepto tradicional de alfabetismo mediático en este modelo Red (Castells, 2003), sino que esta nueva ecología de medios exige el desarrollo de nuevas habilidades, tales como:

“(…) la interpretación crítica de conjuntos (hiper)textuales interactivos, hasta la integración de diferentes tipos de alfabetismo (convergencia de competencias interpretativas vinculadas al cine, la televisión, los videojuegos, etc.) o la capacidad para navegar en las redes digitales, construirse una identidad y on line participar en comunidades virtuales.” (Scolari, 2016, p.7)

Scolari lo denomina “alfabetismo transmedia” y en esta nueva literacidad propone la necesidad de integrar las culturas colaborativas (Jenkins, 2006) y el aprendizaje informal (Sefton-Green, 2006 y 2013) como formadoras de competencias fuera de la escuela y en diálogo con los medios.

2. La gobernanza algorítmica

En la sociedad actual, muchas de nuestras actividades cotidianas están mediadas por algoritmos. Estos sistemas automatizados de recomendación son ampliamente utilizados por las principales plataformas sociales, tales como Google, YouTube, Spotify, Netflix, entre muchas otras. Dichos softwares permiten a estas plataformas resolver problemas computacionales de manera eficiente, entregando predicciones de información relevante para sus usuarios (Bucher, 2018).

El concepto de gobernanza algorítmica es definido como un conjunto de procesos automatizados de recomendación que toman decisiones, pero que tienen efectos sociales (Pasquale, 2015; Kitchin, 2017). De esta manera, los profesionales de las ciencias de la computación y matemáticos son los gestores del desarrollo de nuevas formas de gobierno algorítmicos (Danaher et al., 2017).

Pero, a pesar de que convivimos a diario con dichos algoritmos, para la mayoría de las personas sus acciones resultan opacas (Burrell, 2016). Esta opacidad algorítmica ha sido problematizada por Pasquale (2015) apuntando a los gobiernos y corporaciones que avanzan en modelos de seguimiento permanente y ubicuo de los ciudadanos por medio de algoritmos de caja negra. Este concepto hace referencia a la capacidad que tenemos de observar las entradas y las salidas de un sistema tecnológico, pero el desconocimiento sobre su operación interna, ni qué tan lejos llega la información capturada ni mucho menos las consecuencias que podría tener para nuestras sociedades. De este modo, Pasquale (2015) explica que esta opacidad es justificada, férreamente, por las compañías detrás del concepto de propiedad algorítmica o secreto profesional. Esto ha permitido a las compañías la libertad de recopilar todo tipo de datos de sus clientes. Para estas corporaciones el conocimiento es poder, por lo que insisten en coleccionar cada día más datos de los consumidores de servicios y productos para elaborar detallados perfiles de comportamiento que les permitan mejorar sus estrategias de negocio. Sumado a lo anterior, buscan preservar su espacio de análisis a puertas cerradas fuera del alcance de legislaciones que brinden oportunidades de control a los usuarios sobre sus datos. Por el contrario, existen investigadores que consideran que las características técnicas de los algoritmos les confieren racionalidad, autonomía y objetividad; sin embargo,

en su estudio no analizan el desempeño de los algoritmos en contextos de uso específico (Roberge, 2018).

Pero, a pesar de que para la mayoría de las personas los algoritmos son opacos, sus efectos son observables en diferentes circunstancias de nuestra cotidianidad, siendo capaces de modificar el comportamiento de los ciudadanos (Caplan y Boyd, 2016). Se ha demostrado que la acción algorítmica afecta a nuestras relaciones culturales (Pasquale, 2015; Kitchin, 2017). Es más, los algoritmos están tomando decisiones de manera automática en muchos campos de nuestras vidas, entre ellos, la asignación de créditos, el conteo de votos, el cálculo de impuestos (Kroll, 2015) y en procesos de selección basados en nuestro yo digital que podrían dar paso a discriminación en función de la red en la que participamos (Boyd, Levy y Marwick, 2014). Estos efectos algorítmicos, han impactado también en sistemas financieros o ecológicos, por ejemplo, Karppi y Crawford (2016) presentan la estrecha relación entre las redes sociales, con toda su fragilidad, y el sistema financiero que se alimenta de las conductas de los usuarios. En este caso particular, un tweet generado de manera fraudulenta desde la cuenta de una empresa en particular gatilló una cadena de decisiones que provocaron que el mercado financiero fuera modificado de manera negativa. Esto nos muestra la fuerza de las redes sociales y su influencia en algoritmos automatizados que registran estas plataformas en busca de tendencias.

En esta línea, Caplan y Boyd (2016) describen el rol de los algoritmos en la transformación de la esfera pública desde la perspectiva de los medios, presentando estudios que sustentan esta operación algorítmica. En primer lugar, los algoritmos han influido en las elecciones por medio de la difusión personalizada de mensajes a usuarios indecisos. En segundo lugar, los algoritmos filtran, clasifican y priorizan contenido en base al conocimiento de los usuarios, convirtiéndose en curadores de información para los usuarios. En tercer lugar, los algoritmos han sido operacionalizados por gobiernos para controlar la agenda pública de manera soslayada. Mientras que los *bots* están siendo creados para generar contenido y manipular las decisiones de los usuarios por medio del flujo de información de manera programada. Incluso, los propios algoritmos están siendo capaces de crear noticias con apariencia humana.

La pregunta es ¿Quién entrena el algoritmo? Pues bien, estos dispositivos son contruidos por especialistas en función de las necesidades y reglas de cada área de aplicación, los que intentan validar el cumplimiento de las normas y regulaciones que gobiernan cada uno de los procesos implementados en un producto de software. Pero, este conocimiento experto en la construcción de algoritmos es insuficiente porque no tienen el alcance de validación de la justicia legal respecto de las decisiones tomadas por dichos productos, debido a que las técnicas de evaluación se sustentan en datos históricos registrados por sistemas operados por usuarios que, naturalmente, tienen sesgos en sus decisiones (Kroll, 2015). La producción de modelos computacionales a través de algoritmos de aprendizaje automático - a partir del consumo de datos de comportamiento - están generando sesgos que provienen desde las relaciones sociales que son capturadas por los sistemas informáticos, vale decir, los algoritmos aprenden a partir de los sesgos que capturan de la evidencia histórica de sus propios usuarios (Kroll, 2015; Caplan y Boyd, 2016; Baeza-Yates, 2018). Chander (2016), explica que la operación de los algoritmos en el mundo real, en el cual se producen discriminaciones y sesgos, ocasiona que su aprendizaje se conforme de un conjunto de reglas discriminatorias y sesgadas.

Danaher y sus colegas (2017) desarrollaron un taller de discusión en torno a la influencia de los algoritmos sobre las sociedades, considerando las barreras que impiden una legítima operación de los algoritmos y los procedimientos metodológicos utilizados para los estudios de dichas barreras. A partir de una técnica colaborativa, los participantes del taller establecieron una agenda de temas de investigación clave en esta cuestión:

1. Una de las principales barreras identificadas es la opacidad de los algoritmos, la que impide conocer la forma en la cual opera este dispositivo.
2. Identifican que existe un alto optimismo tecnológico por parte de políticos y tecnólogos, lo que se traduce en un claro sesgo debido al bajo razonamiento crítico de la utilidad de las tecnologías.
3. El pesimismo tecnológico es visto como una barrera que impide construir algoritmos legítimos y efectivos.
4. Identifican un cierto grado de incertidumbre de los efectos de los avances tecnológicos, limitando las oportunidades de desarrollar tecnologías legítimas.

5. Se considera que los profesionales de las tecnologías tienden a carecer de las competencias y conocimientos acerca de las leyes y sistemas de gobierno con los cuales interactúan.
6. Se aprecia una carencia de preparación por parte de los empleados públicos respecto de las tecnologías, lo que se traduce en una barrera de legitimidad, al igual que las capacidades de los legisladores y los sistemas legales.
7. La asimetría existente entre los intereses públicos y privados limitan las posibilidades de desarrollo de productos de software legítimos, debido a las diferencias de intereses entre ambos sectores.
8. Otra barrera en la legitimación de los algoritmos como dispositivos de gobierno es el foco en los individuos más que en la efectividad del gobierno en sí mismo.

En definitiva, los algoritmos se han transformado en objeto de estudio para las ciencias sociales debido a los efectos transparentes en la sociedad (Chander, 2016). En esta línea, existen investigadores que plantean la necesidad de transparentar la operación de los algoritmos para evaluar las decisiones algorítmicas con la finalidad de asegurar los intereses de los ciudadanos y la sociedad en su conjunto (Pasquale, 2015; Kroll, 2015) y en la otra vereda los investigadores que plantean estudiar la acción social de los algoritmos analizando el uso de los usuarios en situaciones concretas (Seaver, 2017).

Boyd y sus colegas (2014) se enfocan, específicamente, en el riesgo de discriminación que representan las decisiones de los algoritmos en temáticas como el acceso a un puesto de trabajo, créditos o conocimiento. Kitchin (2017) considera que los algoritmos son ontogénicos, contingentes y performativos. Desde esta visión, analiza las dimensiones críticas acerca del estudio de los algoritmos y propone su abordaje desde la investigación práctica. Establece que es relevante poner énfasis en el estudio crítico de estos dispositivos culturales, ya que gobiernan gran parte de las decisiones sociales. Los principales desafíos en el estudio de estos dispositivos corresponden al acceso de conocimiento respecto de su proceso de elaboración. Normalmente, se suele ver a estos productos digitales como cajas negras, heterogéneos e incrustados en sistemas complejos. Su estudio se complejiza porque están embebidos en interacciones con otros dispositivos culturales y se desenvuelven en un contexto en función de la contingencia. Es decir, no son dispositivos fijos, sino que

evolucionan a partir de las necesidades del contexto. Por lo que sus efectos pueden ser impredecibles (Kearns y Roth, 2019).

Kitchin (2017) plantea que es necesario sistematizar métodos de estudio empíricos para el análisis de los algoritmos. De esta manera propone acercamientos metodológicos como la revisión exhaustiva del código con el propósito de deconstruir su funcionamiento; la producción reflexiva de código como estrategia de estudio, apoyándose en etnografías que van desde el problema hasta la construcción del código. También propone usar la ingeniería inversa para intentar identificar el comportamiento de algoritmos embebidos en cajas negras. Asimismo, sugiere un acercamiento más alejado del código como la entrevista con profesionales o la etnografía del desempeño profesional en la industria del software. A su vez, propone como una estrategia más compleja analizar los algoritmos en su contexto sociotécnico, para comprender las relaciones que existen entre estos dispositivos culturales. Finalmente, recomienda un acercamiento basado en el comportamiento para analizar el impacto de los algoritmos en sus entornos del mundo real, además de determinar el cumplimiento del propósito para el cual fue construido.

En cambio, Chander (2016) basado en la transparencia de los efectos algorítmicos en la sociedad propone avanzar hacia el análisis de los datos sobre los cuales están aprendiendo los algoritmos con el propósito de descubrir elementos que contribuyan con la discriminación para erradicarlos del entrenamiento de los modelos computacionales y, por tanto, de nuestras vidas.

Un ejemplo concreto del gobierno algorítmico es la decisión de los motores de búsqueda de indexar, o peor aún, dejar de indexar sitios web o documentos, lo que afecta nuestra libertad de acceso a la información de manera oportuna. En este caso los algoritmos de los motores de búsqueda deciden que es lo importante y lo visible (Bucher, 2018). Pero, como hemos revisado, la opacidad de las decisiones del algoritmo forma parte del modelo de negocio de estas compañías. Por su parte, los usuarios parecen no reparar en la parcialidad de estas decisiones. Entonces, al hacer una búsqueda en internet cabe preguntarse, ¿De qué nos estamos perdiendo?

2.2. *La búsqueda de información en internet*

Tal como describimos en el apartado anterior, el acceso a la información indexada en la web es mediado por los algoritmos. Los motores de búsqueda permiten recuperar información indexada en la web a través de su vinculación con las consultas emitidas por los usuarios, según su relevancia (Page et al., 1999), siendo el punto de acceso a la información disponible en la web a través de la elaboración de consultas que sirven de índice para la recuperación, exploración y explotación de la información (Halavais, 2017).

Existe evidencia en la literatura respecto de las dificultades de las personas para iniciar, ejecutar y terminar una búsqueda debido a las decisiones cognitivas asociadas y a la sobrecarga de información. Además, se ha establecido que especialmente las personas más jóvenes tendrían barreras cognitivas propias de su desarrollo, lo que conlleva a un menor esfuerzo durante la búsqueda de información (Duarte Torres et al., 2012; Usta et al., 2014). Zach (2005) establece que el nivel de esfuerzo empleado en la búsqueda depende principalmente de factores externos y de la experiencia de los usuarios. Este menor esfuerzo repercute en las condiciones ambientales asociadas al aprendizaje autónomo y a la paradoja del mayor acceso a la información.

Mientras que la decisión de detener una búsqueda está altamente influenciada por la estimación personal realizada por cada agente respecto de la relevancia de la lista de resultados en la primera página del buscador, la calidad de los términos que son capaces de utilizar, la capacidad de obtener nuevos términos para reformular la consulta y la evaluación personal del esfuerzo necesario para resolver la tarea de búsqueda (Wu y Kelly, 2015). Esto refuerza el bajo esfuerzo de búsqueda evidenciado por los usuarios y que, por tanto, complejiza las condiciones para el aprendizaje autónomo y profundo.

Asimismo, se ha determinado que el proceso de búsqueda se dificulta por:

- a) La diversidad de páginas.
- b) La ambigüedad de las consultas realizadas.
- c) Los bajos niveles de comprensión lectora.

Estos obstáculos se han observado principalmente en niños y jóvenes, los que en su mayoría realizan consultas vagas y redundantes que aumentan la frustración durante la

búsqueda, restringiendo la exploración de los resultados (Foss y Druin, 2014). En palabras de Vilches (2011) “La búsqueda de información en internet es una tarea de aprendizaje, pero también es un estado emocional.” Por lo tanto, si bien los motores de búsqueda proponen una interfaz sencilla basada en la elaboración de consultas en lenguaje natural, esto impone mayor esfuerzo cognitivo y experiencia de parte de los usuarios a la hora de seleccionar los términos que conforman cada una de las frases que son enviadas al sistema de recuperación de información, así como también, al momento de explorar los resultados del motor de búsqueda.

Además, en anteriores investigaciones se han establecido como obstáculos los modelos de personalización de las consultas, tales como: la recomendación de consultas (Duarte Torres et al., 2012) y de detección de propósito del usuario (Sadikov, et al., 2010; Santos et al., 2003) debido a que reducirían nuestro espacio de búsqueda. Estos filtros conforman una burbuja (Pariser, 2017) en la que convergen los registros de idioma, clics (Jiang 2014b) y ubicación geográfica (Jiang, 2014a; Jiang 2014b; Cano-Orón, 2019) del usuario de la plataforma. Restringiendo los resultados, las consultas u ordenando los resultados de búsqueda a partir de esos filtros (Yoganarasimhan, 2020).

En definitiva, los motores de búsqueda son dispositivos tecnológicos que facilitan el acceso a la gran cantidad de información disponible e indexada en internet por medio de una interfaz simple, un simple cuadro de texto, y un conjunto de algoritmos opacos que cruzan los términos del usuario con los documentos. Sin embargo, una serie de variables cognitivas, tecnológicas, sociales y regionales condicionan la experiencia de búsqueda, lo que se traduce en que cada sesión de búsqueda se corresponda con una trayectoria única recorrida por un único usuario en base a sus decisiones, incluso si el objetivo de la tarea es el mismo para todos los usuarios. De este modo, cada vez que un estudiante emite una búsqueda en internet, recibe más información de la que puede gestionar y, en consecuencia, tiende a confiar en la clasificación propuesta por el sistema de búsqueda seleccionando, generalmente, alguna de las tres primeras opciones de la página de resultados (Liu y Zhang, 2019).

En consecuencia, el proceso de búsqueda no es una tarea sencilla, requiere de la destreza y reflexión de los usuarios respecto de sus decisiones y el impacto de éstas sobre su dieta informativa.

- Estrategia de bajada de información mediante *snippet*

Los *snippets* corresponden a pequeñas piezas de información que es posible integrar funcionalmente a módulos de mayor información. En los motores de búsqueda, cada página de resultados (SERP) se compone de una lista ordenada de *snippets*, las que corresponden a un fragmento que aporta información adicional, la cual enriquece el proceso de selección de resultados en la búsqueda, con datos estandarizados como el título, enlace y descripción de la página web.

En una búsqueda de información cada *snippet* opera de forma similar al titular y la bajada de las noticias, ya que es un pequeño fragmento que aporta información relevante para orientar el primer acercamiento de las personas a la información que contiene la página web. Según esto, una *snippet* que ofrece información cercana a la consulta, tiene alta probabilidad de ser seleccionada. Por lo anterior, la decisión de selección de la información también está mediada por la estrategia de bajada de cada *snippet* del sistema de búsqueda.

3. Pensamiento crítico en la búsqueda de información

El pensamiento crítico está inserto en el discurso pedagógico occidental y tiene entre sus máximos logros formar pensadores críticos. Ciertamente, en el discurso educativo existe un amplio consenso respecto de la valoración del pensamiento crítico y su importancia en la formación de los estudiantes, siendo considerado como una de las habilidades esenciales en la educación del siglo XXI (Wilson et al., 2011; Varis y Pérez Tornero, 2012; Reimers y Chung, 2016; Kestin et al., 2017). Por lo anterior, es un tema permanente en la discusión teórica y práctica de la educación. Sobre todo, teniendo en cuenta la actual dificultad para procesar y analizar grandes volúmenes de información. Por consiguiente, es considerado como una herramienta cognitiva clave para combinar conocimiento, habilidades intelectuales y experiencia para lograr eficiencia en el desempeño (Díaz-Larenas et al., 2019).

En definitiva, el pensamiento crítico suele estar presente en la discusión educativa, sobre todo si se trata de mejorar su calidad. Sin embargo, aún no existe consenso sobre cómo abordar su enseñanza. En este punto es posible identificar dos corrientes:

Primero, los que creen que aprender a pensar críticamente debe ser un fin en sí mismo, por lo que es posible enseñar a pensar críticamente de manera general y transversal. En este grupo, autores como Siegel apuestan por la formación de “1. la habilidad de razonar bien (...) y 2. la disposición e inclinación a ser guiado por razones así evaluadas” (2010, p.142). Para autores como Standish y Thoilliez (2018) para alcanzar un pensamiento auténticamente crítico se requiere de la transmisión de contenidos disciplinares ricos en ideas e insertos en una cultura pedagógica que fomente el diálogo y el debate constante.

Segundo, aquellos que por el contrario creen que la única opción de aprender a pensar críticamente es mediante el contenido disciplinar. Este último argumento se sostiene en la educación liberal, para quienes “pensar bien” sobre algo (entendiendo pensar bien como lo plantea Ennis) requiere atender a los criterios de la disciplina. De este modo, no es posible formar competencias críticas sin la mediación de contenidos específicos (McPeck, 2016).

A pesar de esta discusión, en la educación superior existe consenso respecto de la necesidad de promover el pensamiento crítico como parte integral de la formación de los estudiantes como ya lo declaraba el informe Delors (1995) “Mientras la sociedad de la información se desarrolla y multiplica las posibilidades de acceso a los datos y a los hechos, la educación debe permitir que todos puedan aprovechar esta información, recabarla, seleccionarla, ordenarla, manejarla y utilizarla” (p.18).

Asimismo, el concepto de alfabetización mediática plantea que es necesario proveer a las personas de las competencias críticas y herramientas necesarias para desenvolverse como ciudadanos autónomos capaces de realizar un consumo crítico de los medios y la información disponible (Buckingham, 2007 y 2008; Gozávez-Pérez y Aguaded-Gómez, 2012).

En esta misma línea, la iniciativa de las Naciones Unidas “Transformando nuestro mundo: la Agenda 2030 para el Desarrollo” (Kestin et al., 2017), en sus 17 objetivos de desarrollo sostenible (ODS) consideran que una educación de calidad incluye la formación del pensamiento crítico como parte de las habilidades requeridas para empoderar y movilizar a los estudiantes para transformarlos en ejecutores efectivos de ODS. De esta forma, todos los estudiantes egresados de la educación superior requieren de:

“Habilidades transversales y competencias clave que son relevantes para abordar todos los ODS: pensamiento sistémico, pensamiento crítico, autoconciencia, resolución integral de problemas y capacidad de anticipación, competencias normativas, estratégicas y de colaboración; creatividad, espíritu emprendedor, curiosidad y habilidades de aprendizaje, pensamiento creativo, responsabilidad social, competencias de asociación y sentirse cómodo en entornos interdisciplinarios” (p.12).

Como detallamos en los antecedentes del desarrollo de la educomunicación, el currículum AMI para profesores de la UNESCO (Wilson et al., 2011) entrega los elementos clave para la alfabetización mediática, entre los que considera la evaluación crítica del contenido de los medios. Sin embargo, a pesar de la constante promoción de la necesidad de formar a los estudiantes en competencias AMI, la inclusión de una perspectiva crítica en la práctica pedagógica no ha tenido el avance necesario (González-Osorio, 2021; Mesquita-Romero et al., 2022), tal como lo constatamos durante el confinamiento pandémico.

3.1 Motivación: la clave del pensamiento crítico

Desde la perspectiva educativa la motivación es un aspecto clave para el aprendizaje de los estudiantes y ha sido modelada basándose en la orientación a la tarea (Eccles y Wigfield, 2002) y por las motivaciones personales del estudiante para aprender (Valenzuela et al., 2015). En esta investigación nos basamos en la propuesta de Valenzuela (Valenzuela y Nieto, 2008), según la cual la motivación escolar corresponde a la activación de los recursos cognitivos usados para aprender lo propuesto por la escuela como aprendizaje. Según esto, la motivación escolar es una combinación entre realizar la tarea y motivarse por la actividad de aprendizaje. Si bien el autor hace referencia al contexto escolar, su definición de motivación al pensamiento crítico es aplicable al contexto universitario y, específicamente, a esta investigación.

No obstante, es necesario relevar que la activación del pensamiento crítico depende de la motivación del estudiante a pensar críticamente, cuya decisión está sujeta a la valoración del estudiante respecto de la importancia, utilidad, costo e interés personal (Saiz y Rivas, 2008; Valenzuela y Nieto, 2008). De acuerdo con lo anterior, para alcanzar una auténtica

reflexión crítica del estudiante es necesario elaborar tareas de aprendizaje que sean relevantes, útiles, motivadoras y cuyo desarrollo sea posible de alcanzar por los aprendices.

En esta investigación se problematiza respecto de la posibilidad que tienen los estudiantes de actuar críticamente en la selección de información, considerando la saturación de información en internet. ¿Es posible ser crítico durante una búsqueda de información? ¿La búsqueda de información es un proceso consciente o automático?

4. Sesgos cognitivos

La larga data de la investigación respecto de la capacidad humana de pensar racionalmente nos ha mostrado que las personas no tomamos buenas decisiones. Es más, las decisiones que tomamos a diario están lejos de ser racionales, normalmente, nuestras decisiones respecto del cálculo y la probabilidad son contrarias a las leyes de la probabilidad (Cortada, 2008). Simon (1957) la denominó teoría de racionalidad limitada, según esta teoría nuestras decisiones racionales están limitadas respecto de las restricciones inherentes de nuestra búsqueda y nuestra capacidad de cálculo. Basados en esta propuesta de los límites del pensamiento humano, Tversky y Kahneman (1983) plantean su teoría de sistemas de pensamiento, según la cual las personas no somos seres irracionales, sino que contamos dos formas de pensamiento diferentes.

De acuerdo con esta teoría de pensamiento dual, los seres humanos tenemos dos sistemas de pensamiento: el sistema 1 y el sistema 2 (Kahneman, 2012). El sistema 1 corresponde al sistema más antiguo, que nos acompaña desde la evolución y cuya toma de decisiones está basada en la intuición. Este sistema lo compartimos con los demás animales y es el que ha permitido la sobrevivencia de la especie humana, ya que su eficiente y rápido proceso nos facilita tomar decisiones sin un gran esfuerzo cognitivo (Matute, 2019). El sistema 1 se activa de forma automática y se conduce mediante el hábito, esta característica permite que lo utilicemos la mayor parte del tiempo. Lo anterior, implica que, por un lado, nos permite tomar decisiones en estados de incertidumbre y ser eficientes en multitareas, ya que trabaja en paralelo; pero, por otro lado, su automatismo nos dificulta controlarlo o modificarlo (Cortada, 2008).

Por el contrario, el sistema 2 corresponde al pensamiento racional y es el más reciente en la evolución humana. Se caracteriza por ser lento, controlado y por el gran esfuerzo cognitivo que exige. Además, tal como señala Matute (2019) el sistema 2 no se activa por defecto al nacer como ocurre con el sistema 1, sino que se configura con el entrenamiento constante y prolongado en el tiempo. De las diferencias de los sistemas 1 y 2 surgen dos estrategias para solucionar problemas: las decisiones racionales, es decir, los algoritmos, las reglas y estrategias que son siempre correctos; y las decisiones basadas en heurísticos o juicios intuitivos, los cuales podrían resultar correctos o erróneos (Cortada, 2008). Desde esta teoría, la activación de la motivación a pensar críticamente se adscribe al sistema 2 de pensamiento, lo que conlleva un mayor esfuerzo cognitivo de la persona.

En consecuencia, esta teoría nos invita a comprender que cohabitamos con nuestros sesgos y los de los demás. No obstante, en las situaciones que así lo requieren es necesario accionar el sistema 2 y evitar ser conducidos por sesgos cognitivos.

4.1 Sesgos cognitivos en la búsqueda de información

La búsqueda de información en internet está sujeta a las reglas de la web. De este modo, recuperar información requiere de la mediación de un algoritmo de búsqueda. En internet existen diferentes motores de búsqueda y cada uno de ellos cuenta con un algoritmo de búsqueda, con sus propios criterios de indexación y recuperación de información. De esta forma, no ha sido posible determinar si, ante una consulta idéntica, es probable obtener la misma fuente de información o el mismo orden de relevancia en diferentes motores de búsqueda (Yagci, et al., 2022).

En la literatura es posible encontrar diversos autores que exponen la ocurrencia de sesgos cognitivos en internet (Kroll, 2015; Caplan y Boyd, 2016; Baeza-Yates, 2018). Baeza-Yates (2018), describe un círculo vicioso en el que se observan sesgos en la búsqueda de información. De este modo, los sesgos cognitivos propios de nuestra condición humana han sido traspasados a internet. Primero, por la intervención de los programadores en la creación de los algoritmos. Segundo, por las personas creadoras de contenidos en la web. Tercero, por la interacción entre las personas y los diferentes contenidos de la web.

En este sentido, cada una de las personas que participa de la web contribuye y, a la vez, queda expuesta a recibir información parcial o sesgada. Cabe mencionar, que no atribuimos una voluntad de sesgar la información, sino que, dada la velocidad y volumen de información las decisiones cognitivas se tornan cada vez más complejas.

Por lo anterior, en esta investigación, se analizan los sesgos cognitivos en las búsquedas de información en internet como producto de un procesamiento de información automático no intencionado. Debido a las limitaciones cognitivas de las personas para procesar tareas complejas en ambientes de incertidumbre debido a la rapidez y sobrecarga de información, en los cuales los heurísticos nos permiten tomar decisiones rápidas para aceptar o rechazar la información recibida. Sin embargo, si estos se transforman en sesgos nos conducen a errores.

OBJETIVO DE LA INVESTIGACIÓN

De la revisión de la literatura se infiere la necesidad de estudiar el uso de los motores de búsqueda en la educación superior como dispositivos habilitadores del aprendizaje autónomo y profundo. Teniendo en cuenta que los modelos educativos vigentes en las principales instituciones de educación superior basan su práctica pedagógica en la implementación de metodologías activas, donde los estudiantes son los principales responsables de sus aprendizajes. Asimismo, los docentes y los estudiantes utilizan estos sistemas para obtener información actualizada para construir conocimiento situado y pertinente. Los sistemas de recuperación de información brindan la valiosa oportunidad de acceder de manera rápida y efectiva a la información. Sin embargo, tal como se menciona en anteriores estudios, se ha establecido que para un aprovechamiento virtuoso de esta tecnología es decisivo contar con las competencias críticas y técnicas para localizar y seleccionar la información fiable (Druin et al., 2009; Qureshi et al., 2015; Şendurur & Yildirim, 2015). Asimismo, en recientes investigaciones se ha establecido que en la educación superior los estudiantes no disponen de las competencias necesarias para beneficiarse de esta tecnología. Normalmente, aprenden por sus propios medios mediante la práctica, sin una guía efectiva (Pereira et al., 2019). En definitiva, la búsqueda de información

es una acción cotidiana en actividades educativas donde la mayoría de los estudiantes sabe cómo elaborar una consulta en el buscador, pero para una búsqueda eficiente se requieren habilidades críticas para discriminar la información disponible.

Actualmente, estos algoritmos son capaces de inducir modelos generalizados de comportamientos o fenómenos, a partir de grandes cantidades de datos. Existe una corriente de investigación preocupada por las implicancias de los modelos computacionales inferidos desde datos masivos sobre la vida cotidiana de las personas, debido a las modificaciones del comportamiento (Chander, 2016; Boyd et al., 2014). Lo anterior, ha incrementado su tratamiento en la literatura, traspasando las fronteras de la informática para ser un objeto de interés de las ciencias sociales, principalmente en lo relativo al uso, manipulación y privacidad de la información (Burrell, 2016, p2.). Además, las decisiones algorítmicas han sido definidas como opacas (Pasquale, 2015) porque no es posible conocer con exactitud las decisiones del modelo inferido desde el algoritmo. Por lo tanto, se ignoran las reglas algorítmicas aplicadas sobre los datos y, en particular, las que seleccionan nuestros resultados en el motor de búsqueda.

La creciente saturación de información ha dado paso a la personalización de los resultados por medio de nuevos algoritmos. Esto ha complejizado la interacción entre las personas y la información, ya que han limitado el acceso a información diversa que gatille la activación de las habilidades críticas. Por lo tanto, el aumento de la saturación en los resultados de búsqueda sumado a una baja motivación al pensamiento crítico aumentaría la operación de los sesgos cognitivos sobre la estrategia de exploración, selección y análisis de la información.

En consecuencia, la mediación algorítmica acotó los resultados y, por tanto, la dieta informativa, de manera transversal a todos los usuarios, limitando el acceso a la información disponible en función de las decisiones de los usuarios, con consecuencias epistémicas negativas al consumo de información en línea. Por lo tanto, la oferta de información saturada y filtrada restringen el acceso a información nueva, lo que reduciría la puesta en funcionamiento de las habilidades del pensamiento crítico. En cualquier caso, todos los estudiantes se ven enfrentados a las mismas decisiones algorítmicas, pero los resultados dependen de la interacción entre sus decisiones y los algoritmos.

En la revisión de la literatura se encontraron estudios recientes que analizan los sesgos en el comportamiento de los usuarios, a partir de la manipulación de los resultados de búsqueda (Novin & Meyers, 2017), las decisiones de los usuarios (Ghenai et al., 2020), el tiempo invertido en la búsqueda (Pothirattanachaikul, et al., 2020) y en la evaluación de los resultados (Liu, 2019). Estos estudios confirman que a partir de la caracterización del comportamiento de búsqueda de las personas es posible identificar la operación de los sesgos cognitivos inducidos por los motores de búsqueda. En cuanto a la influencia de la saturación en el comportamiento de búsqueda se ha determinado que los estudiantes que reciben resultados no saturados visitan más páginas y dedican más tiempo a la exploración de la web que quienes recibieron resultados saturados (Kelly & Azzopardi, 2015). Sin embargo, no se encontraron estudios que establezcan la relación de la saturación de información sobre la operación de los sesgos cognitivos derivados de las decisiones de los usuarios en contextos naturalistas.

En esta investigación buscamos indagar en el uso del motor de búsqueda como la principal plataforma de búsqueda de información en internet. Entendiendo que, a pesar de su difundido uso, las personas no conocemos los criterios de indexación de los algoritmos de los buscadores ni cuestionamos el alcance de la influencia del sistema de búsqueda en lo que denominamos dieta informativa personalizada.

Primero, analizamos las implicancias de la mediación algorítmica en la elaboración de nuestra dieta informativa, configurada a partir del registro de nuestros patrones de comportamiento en la web. Posteriormente, centramos nuestro interés en estudiar por medio de dos acercamientos empíricos los patrones de comportamiento de los sistemas de búsqueda de estudiantes universitarios chilenos. De esta forma los diseños cuasi experimental y experimental nos permiten obtener evidencias para caracterizar las decisiones cognitivas de los estudiantes durante sesiones de búsqueda de información para resolver tareas académicas en un contexto que se asemeje al natural. Para profundizar la interpretación de los datos recogidos de la sesión experimental se realizan entrevistas semiestructuradas que permiten enriquecer el conocimiento del fenómeno en estudio.

1. Objetivo general

Evaluar la influencia de la interacción entre el nivel de saturación de información en los resultados de búsqueda en la web y la motivación al pensamiento crítico sobre la operación de los sesgos cognitivos durante la exploración, selección y análisis crítico de fuentes de información en la búsqueda de información en la web por parte de estudiantes universitarios.

2. Objetivos

1. Caracterizar las decisiones cognitivas (sesgos, calidad de fuentes y mediatización) por medio de un estudio de contenido del comportamiento de búsqueda de los estudiantes que participan de la investigación.

2. Establecer la influencia de la saturación de información en los resultados de búsqueda en la web sobre la operación de sesgos cognitivos por medio de un estudio de contenido entre las fuentes seleccionadas y los resúmenes elaborados por los estudiantes participantes, considerando la interacción con la motivación al pensamiento crítico.

3. Establecer la coherencia entre la operación de sesgos cognitivos identificados en el comportamiento de búsqueda de información en la web y la reflexión reportada por los participantes acerca de las decisiones de inclusión y exclusión de documentos.

DISEÑO METODOLÓGICO

Para alcanzar el propósito de esta investigación se eligió un diseño mixto secuencial (Creswell, 2013). En la primera etapa cuantitativa los estudiantes respondieron un test de motivación al pensamiento crítico para determinar su nivel de motivación; posteriormente, realizaron búsquedas de información en internet en un ambiente naturalista. Los resultados experimentales fueron profundizados a través de una segunda etapa cualitativa que consistió en el desarrollo de entrevistas semi estructuradas para ahondar el análisis de las decisiones de los participantes durante el proceso de búsqueda de información en internet. Se espera establecer que un alto nivel de motivación al pensamiento crítico reduce la operación de los sesgos cognitivos, independiente del nivel de saturación de la información, estableciendo las

estrategias de búsquedas críticas de información. Mientras que, frente a mayores niveles de saturación de información, se observa una mayor operación de sesgos cognitivos derivados de la frustración o desgano.

Para esta investigación se diseñaron dos acercamientos experimentales. Ambos diseños metodológicos fueron previamente aprobados por el Comité de Ética y Bioética de la Universidad Austral de Chile.

1. Primer estudio cuasi- experimental

El diseño del primer estudio tiene un enfoque cuasi experimental. En este cuasiexperimento se analiza el comportamiento de los participantes durante una sesión de búsqueda de información utilizando la plataforma GoNSA2. Se pretende caracterizar las decisiones de búsqueda de información de los estudiantes universitarios de primer año mientras resuelven una serie de problemas planteados y validados en la literatura. El estudio es de carácter naturalista, donde los participantes emiten consultas, revisan documentos, seleccionan información y responden al cuestionamiento planteado tal como lo harían de manera natural en sus tareas cotidianas. Este estudio busca evaluar la oportunidad de los estudiantes de acceder a fuentes de calidad a partir de una búsqueda de información.

1.1 Muestra

El universo de estudio corresponde a estudiantes universitarios. Se utilizó una muestra no probabilística por conveniencia (Flick, 2007). El criterio para seleccionar las unidades de análisis es el nivel de formación educativa de los sujetos, por lo que se selecciona a los estudiantes de primer año, ya que tienen un avance similar en la malla curricular de su carrera.

1.2 Estrategia de recolección de datos

La recolección de los datos de búsqueda se realizó de manera automática en la plataforma GoNSA2 (Olivares-Rodríguez et al., 2018) cuyas funcionalidades se describen al final de este capítulo en el apartado 3.

1.3 Análisis de datos

El análisis de las trazas de búsqueda se realizó manualmente a partir de los datos recopilados en la plataforma GoNSA2. Primero, se categorizaron las fuentes de la primera página de resultados de cada una de las tareas. A continuación, se revisaron los procesos editoriales declarados por las webs de la primera página de la lista de resultados para cada una de las tareas propuestas.

1.3.1 RankFlow

Para el análisis de la fluctuación de las páginas webs se utilizó la métrica RankFlow (Rieder, 2016). Rankflow es una herramienta de visualización en línea (<https://labs.polsys.net/tools/rankflow/>) que permite analizar los cambios en listas clasificadas, inclusive, valoradas a lo largo del tiempo. Los cálculos de RankFlow están basados en el trabajo de William Webber y sus colegas (2010). La herramienta permite utilizar datos de una colección de elementos clasificados o combinaciones de elementos y un valor. Esta herramienta también permite calcular la métrica de la distancia sesgada por el rango (RBD) para cuantificar los cambios de un tramo a otro. Cuanto mayor sea el valor de la RBD, mayor será el cambio. El parámetro "RBD p" (cuyo valor está entre 0,01 y 0,99) permite determinar el grado de "ponderación" del cálculo. Con un p pequeño, los cambios en la parte superior de las listas se ponderan más fuertemente, con p acercándose a 1 todos los cambios se tratan igual.

1.4. Tareas de búsqueda

Para el primer acercamiento experimental se utilizaron dos tareas factuales y dos tareas exploratorias previamente validadas en la literatura (Arguello, et al., 2012; Kules y Shneiderman, 2008). Las tareas se muestran en la siguiente tabla:

Tabla 1: Tareas de búsqueda en internet cuasiexperimento

Consigna	Tipo	Tiempo
Señalar el nombre del punto más profundo del mar	factual	5 minutos
Determinar el número de personas con VIH en Chile	factual	5 minutos
Determinar las componentes que respalden un plan de combate del crimen en Chile	exploratoria	15 minutos
Identificar las etapas necesarias para la construcción de un automóvil	exploratoria	15 minutos

Fuente: Elaboración propia

1.5 Protocolo de investigación

El estudio cuasi experimental se realizó de manera presencial. Para la conducción de la investigación se realizaron dos sesiones con los participantes.

Primera sesión: De socialización de los fines de la investigación, los estudiantes recibieron de manera física el documento de consentimiento informado (C.I), por lo que durante esta sesión se recopilaron los C.I firmados por los participantes.

Segunda sesión: Los estudiantes que aceptaron participar de la investigación recibieron en sus correos institucionales las credenciales de acceso a GoNSA2. La comunicación con los estudiantes se realizó de forma presencial en su sala de clases. En primer lugar, se validó que los participantes accedieran de manera correcta a la plataforma y, a continuación, se presentaron las funcionalidades técnicas de la plataforma. Una vez que ingresan a la plataforma, los estudiantes reciben las indicaciones para realizar las tareas de búsqueda y disponen de 1 hora para resolver las cuatro tareas planteadas.

2. Segundo estudio experimental

El segundo acercamiento al objeto de estudio se realizó mediante el diseño de un estudio experimental que permitiera caracterizar la operación de sesgos cognitivos considerando la influencia de dos variables independientes: la saturación de información en interacción con la variable motivación al pensamiento crítico. En el diseño de este experimento definieron las siguientes variables:

2.1. Variables

2.1.2 Variables independientes

Motivación al pensamiento crítico: Al iniciar una búsqueda se ponen en juego las habilidades de pensamiento crítico que permiten definir el problema de información, usar las estrategias pertinentes, localizarlas y relevarlas (Eisenberg y Berkowitz, 2000). Esta investigación se adhiere al concepto de pensamiento crítico que involucra las habilidades cognitivas y, especialmente, las disposiciones del sujeto (Ennis, 1985, 2011; Facione, 1990), las cuales representan el componente emocional del ser humano, entendidas como cualidades esenciales para alcanzar una reflexión crítica en los estudiantes sobre la información, sobre todo, considerando que las habilidades de acceso no son garantía de éxito en un ambiente saturado de información.

Saturación de información: Corresponde a la sobrecarga de información que reduce la capacidad de tomar decisiones efectivas por parte de los usuarios (Faulkner y Trotter 2017; Roetzel, 2019; de Jong, Bok, Kremer y van der Vleuten, 2019). En particular, durante la búsqueda de información esta sobrecarga se evidencia en la cantidad de resultados por página o en la cantidad total de resultados (Kelly y Azzopardi, 2015).

2.1.3 Variable dependiente

Operación de sesgos cognitivos: Los sesgos cognitivos corresponden a formas no razonadas en las que se toman decisiones, de bajo nivel de esfuerzo cognitivo, que pueden desencadenar la toma de decisiones parciales o erradas (Tversky y Kahneman, 1983). Así, los sesgos operan como filtros para obtención de una respuesta rápida que podría ser eficaz en algunos contextos pero que, a la vez, conduce a errores. En esta tesis, se analizarán los sesgos

cognitivos como producto de un procesamiento de información automático no intencionado, basándonos en las limitaciones cognitivas de los seres humanos para procesar tareas complejas en ambientes sobrecargados de información. Por lo que se analizará el contenido de las evidencias de búsqueda de información para inferir la emergencia de sesgos cognitivos presentes en la web.

2.1.4 Variables extrañas

Área disciplinar: Corresponde a la disciplina de la población de estudio.

Algoritmos de personalización: Conjunto de instrucciones que procesan entradas para generar una salida. Tal como se detalló en el apartado gobernanza algorítmica, estos algoritmos de perfilado son capaces de inducir modelos generalizados de comportamientos o fenómenos, a partir de grandes cantidades de datos históricos de sus usuarios (Kroll, 2015; Caplan y Boyd, 2016; Baeza-Yates, 2018).

2.2 Muestra

El universo de estudio corresponde a estudiantes universitarios. Se utilizó una muestra no probabilística por conveniencia (Flick, 2007). El criterio para seleccionar las unidades de análisis es el nivel de formación educativa de los sujetos, por lo que se selecciona a los estudiantes de primer año, ya que tienen un avance similar en la malla curricular de su carrera. La selección de la muestra corresponde a un proceso aleatorio de la población, mientras que la construcción de los grupos experimentales obedece a un proceso estratificado aleatorio en función de las variables de 1) motivación al pensamiento crítico y 2) área disciplinar, lo que permite distribuir la muestra de manera homogénea en los grupos y, con ello, controlar ambas variables.

2.3 Estrategia de recolección de datos

La recolección de datos corresponde a la captura del nivel de motivación al pensamiento crítico, la extracción de las trazas de búsqueda de información para analizar la operación de los sesgos cognitivos y las entrevistas semi estructuradas a los participantes.

Para determinar la motivación al pensamiento crítico se utilizará la Escala Motivacional de pensamiento Crítico, EMPC (Valenzuela, 2008; Valenzuela et al., 2017), la

cual estudia la relación entre la motivación y sus componentes y el pensamiento crítico. La EMPC se compone de 20 ítems tipo Likert con puntuaciones de 1 a 6, en la cual los sujetos expresan su grado de acuerdo o desacuerdo de una serie de afirmaciones respecto de las expectativas que tienen de pensar críticamente y, de afirmaciones valorativas respecto de la importancia y utilidad percibida hacia el pensamiento crítico, el costo que asumirían por hacerlo y su interés personal en pensar de forma rigurosa. Así, los ítems de la EMPC son: Expectativa, Valor, Utilidad, Interés y Costo, los que se distinguen factorialmente en 5 subescalas en 1er y 2do orden ($KMO = ,887$ y $,902$), correspondientes a los constructos propuestos por el modelo de Eccles y Wigfield (2002) con altos niveles de confiabilidad. Para el ítem Expectativa ($\alpha = ,774$), Importancia ($\alpha = ,770$), Costo ($\alpha = ,775$), Utilidad ($\alpha = ,790$) e Interés ($\alpha = ,724$).

Para determinar la operación de los sesgos cognitivos en la búsqueda de información se analizarán las estrategias de exploración, selección y análisis de información de los participantes a partir de las trazas de búsqueda de información, las cuales se extraen desde la plataforma de búsqueda de información desarrollada en el Instituto de Informática de la UACH, denominada GoNSA2 (Olivares-Rodríguez et al., 2017). Esta interfaz permite capturar, en un ambiente realista, las estrategias de búsqueda por parte de los participantes.

Para la identificación de la operación de los sesgos cognitivos se propone observar la operación de los sesgos cognitivos en: 1) las consultas emitidas por los estudiantes, 2) el contenido de los documentos seleccionados en cada una de las sesiones de búsqueda de información en la web, 3) el contenido de los resúmenes elaborados y 4) las entrevistas semi estructuradas.

Previo al análisis de las trazas de búsqueda obtenidas desde la plataforma GONSA2 se elaboraron las reglas de análisis para observar la ocurrencia de sesgos cognitivos. Estas reglas sirvieron de base para observar la operación de los sesgos durante el proceso de búsqueda. Sin embargo, durante el análisis se identificaron y caracterizaron otros sesgos cognitivos no previstos.

Finalmente, para determinar las creencias, prácticas y los criterios de selección y exclusión de documentos se ejecutaron entrevistas semi-estructuradas (Flick, 2007). Esta

técnica aportó “(...) riqueza informativa de las palabras y las interpretaciones de los entrevistados (...)” (Sautu, et al., 2005, p. 49), permitiendo hacer preguntas que permitieron clarificar las decisiones tomadas por los estudiantes durante la búsqueda y elaboración de sus resúmenes. Los entrevistados fueron voluntarios, de ambos géneros y carreras y, en relación con la variable independiente de la motivación al pensamiento crítico en interacción con el nivel de saturación de la información.

2.4 Análisis de datos

Para caracterizar la operación de sesgos cognitivos (o) por medio de un estudio de contenido del comportamiento de búsqueda (consultas, fuentes seleccionadas, resúmenes) de los estudiantes que participan del experimento, se propone un estudio sobre la diversidad interna de los contenidos presentes en las consultas, en los documentos seleccionados y en los resúmenes elaborados. Asimismo, se propone un análisis cuantitativo de indicadores vinculados a los diferentes sesgos cognitivos presentes en las trazas del comportamiento de búsqueda de información.

Para evaluar la influencia de la saturación de información (s) en los resultados de búsqueda en la web sobre la operación de sesgos cognitivos (o) por medio de un estudio de contenido entre las fuentes disponibles y las seleccionadas por los estudiantes participantes, se propone un análisis de contenido para determinar el contraste entre la información disponible en los resultados de búsqueda y la información presente en los documentos seleccionados, considerando el nivel de motivación de pensamiento crítico de los participantes. El análisis cuantitativo de las consultas y documentos permite determinar el nivel de operación de los sesgos cognitivos y, por tanto, es posible realizar un análisis estadístico (ANOVA) entre las variables en estudio. Asimismo, se propone un análisis cuantitativo de indicadores vinculados a los sesgos cognitivos en los documentos seleccionados, en función de los niveles de saturación de información presentados en el experimento.

Para determinar la influencia de la saturación de información (s) en los resultados de búsqueda en la web sobre la operación de sesgos cognitivos (o) a partir del análisis de contenido de los resúmenes elaborados por los estudiantes respecto a los documentos

seleccionados, se propone un análisis de contraste de contenido entre los argumentos elaborados en los resúmenes, los documentos seleccionados y los documentos disponibles en los resultados de búsqueda, considerando el nivel de motivación de pensamiento crítico de los participantes. El análisis cuantitativo de los resúmenes permite determinar el nivel de operación de los sesgos cognitivos y, por tanto, es posible realizar un análisis estadístico (ANOVA) entre las variables en estudio. Asimismo, se propone un análisis cuantitativo de indicadores vinculados a los sesgos cognitivos en los argumentos, en función de los niveles de saturación de información presentados en el experimento.

Para evaluar la coherencia entre la operación de sesgos cognitivos identificados en el comportamiento de búsqueda de información en la web y la reflexión reportada por los participantes acerca de las decisiones de inclusión y exclusión de documentos, se propone un estudio cualitativo a partir de una entrevista semiestructurada. Esto permite indagar en las decisiones de los participantes respecto a la selección y análisis de la información para, con ello, contrastar esta información con su comportamiento de búsqueda en la web. Se eligió este tipo de entrevista porque permite la selección muestral de los entrevistados y la codificación a posteriori de las respuestas, lo que permitió incluir las preguntas abiertas y aquellas no previstas (Gaitán y Piñuel, 2010).

2.5 Tareas de búsqueda

Para el segundo acercamiento experimental se elaboraron tareas propias, de acuerdo con los lineamientos de la taxonomía bidimensional de Anderson y Krathwohl (Urgo et al., 2019) incluyendo temáticas con alto nivel de mediatización que permitieran nivelar el conocimiento del dominio para todos los participantes. Además, se consideró necesario reducir la sesión experimental al desarrollo de 3 tareas para reducir el abandono observado en la última tarea durante el primer cuasiexperimento. De esta manera, se estableció evaluar el paquete de tareas mediante un comité de expertos. Con esta finalidad se elaboraron siete tareas (tabla 2), las que fueron sometidas a la evaluación de tres académicos expertos en: Educación y Tecnología, Neurodidáctica y Lingüística. Entre las tareas con mayor acuerdo se seleccionó una tarea factual y dos exploratorias, tal como se observa en la Ilustración 1.

Tabla 2: Propuesta de tareas de búsqueda en internet

Consigna tarea	Tipo	Tiempo
Mencione la ley chilena que tipifica el femicidio	factual	5 minutos
Mencione la cantidad de personas fallecidas por COVID-19 en Chile el 2020	factual	5 minutos
Compare el plan paso a paso de Chile con el plan de prevención del COVID 19 de Colombia	exploratoria	15 minutos
Resuma las características de una noticia falsa	exploratoria	15 minutos
Compare la eficacia de las vacunas Sinovac, Pfizer y AstraZeneca	exploratoria	15 minutos
Resuma las razones del conflicto de la plaza Baquedano	exploratoria	15 minutos
Resuma las razones del cambio de fecha de las elecciones "Abril 2021"	exploratoria	15 minutos

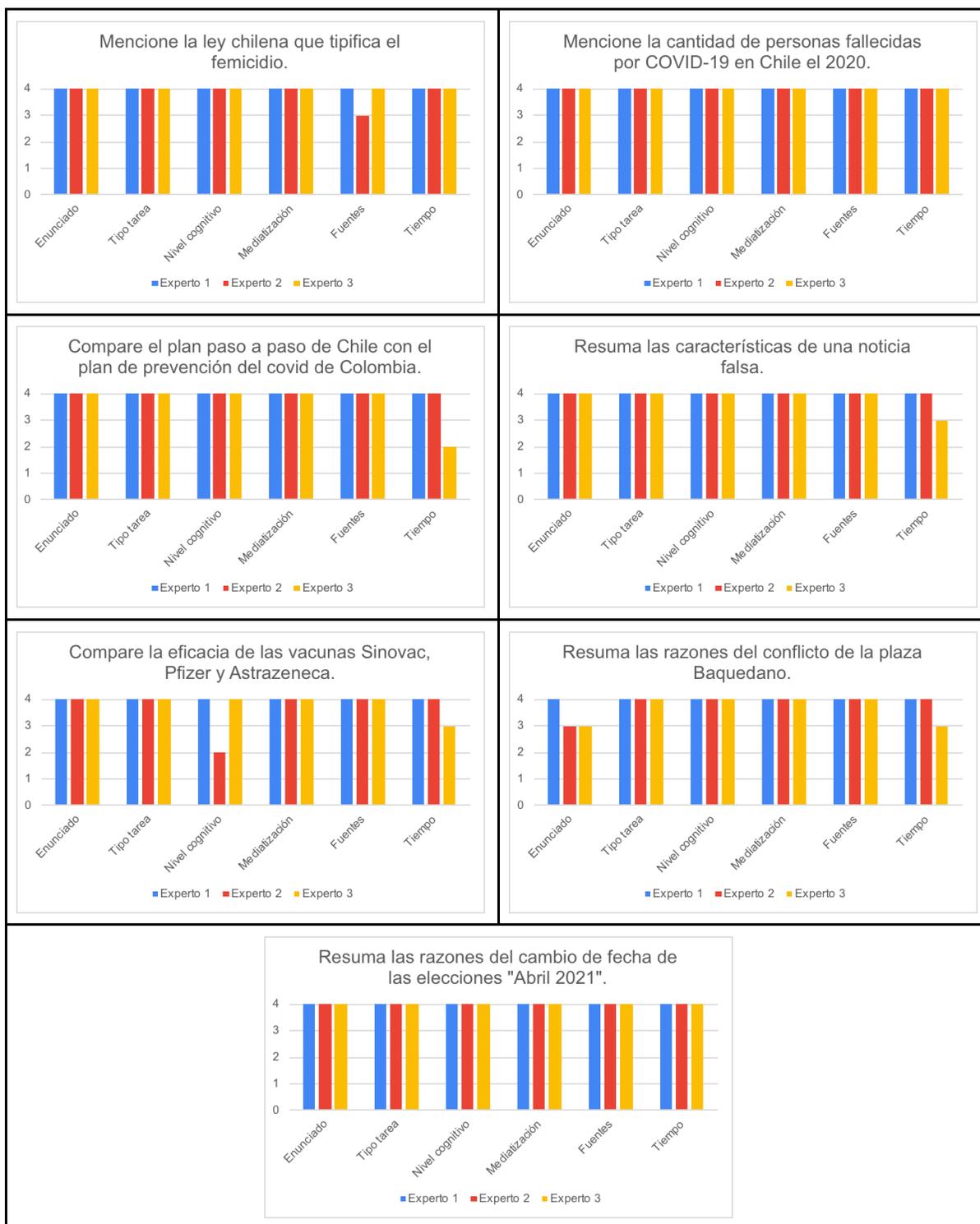
Fuente: Elaboración propia

La evaluación de las tareas se realizó mediante un formulario Google Forms ® en el cual se presentaron cada una de las tareas y se solicitaba a los expertos evaluar las siguientes dimensiones:

Tabla 3: Cuestionario de evaluación del juicio de expertos

Dimensión	Preguntas
Comprensión del enunciado	¿Los términos utilizados en el enunciado de la tarea son adecuados para facilitar su comprensión por estudiantes universitarios de distintas disciplinas? Argumente su respuesta.
Tipo de tarea	¿La tarea corresponde a una tarea de tipo factual de búsqueda de datos concretos? Argumente su respuesta. ¿La tarea corresponde a una tarea exploratoria? Argumente su respuesta.
Nivel cognitivo de la tarea	¿La tarea corresponde al nivel cognitivo recordar? Argumente su respuesta. ¿La tarea corresponde al nivel cognitivo comprender? Argumente su respuesta.
Mediatización del contexto de la tarea	¿El contexto de la tarea corresponde a un tema presente en la agenda mediática del país y, por tanto, de dominio público? Argumente su respuesta.
Fuentes de información	Al elaborar una consulta en internet ¿Existe diversidad de fuentes de información respecto del tema de la tarea? Argumente su respuesta.
Tiempo de desarrollo	¿El tiempo asignado es adecuado para resolver la tarea?

Ilustración 1: Resultados evaluación expertos



2.6 Protocolo de investigación

El segundo experimento se realizó en tres sesiones virtuales/síncronas usando la plataforma Zoom®, de acuerdo con el reglamento sanitario de la universidad Austral de Chile durante la pandemia COVID-19.

Primera sesión: De socialización de los fines de la investigación, previamente los estudiantes recibieron en sus correos institucionales el documento de consentimiento informado (C.I), por lo que durante esta sesión se recopilaron los C.I de los participantes. A continuación, se realizó el test de EMPC solo a los estudiantes que enviaron su C.I, utilizando un formulario Google®.

Segunda sesión: Los estudiantes que aceptaron participar de la investigación recibieron en sus correos institucionales las credenciales de acceso a GoNSA2. La comunicación con los estudiantes se realizó de forma síncrona mediante la plataforma Zoom®, lo que permitió apoyar el acceso y presentar las funcionalidades técnicas de la plataforma. Una vez que ingresan a la plataforma, los estudiantes reciben las indicaciones para realizar las tareas de búsqueda y disponen de 45 minutos para resolver las tres tareas de búsqueda.

Tercera sesión: Participaron 14 estudiantes de ambas carreras, los que fueron seleccionados de manera aleatoria. Dichos estudiantes fueron invitados a participar, voluntariamente, mediante un correo electrónico para agendar un horario. Las entrevistas se realizaron de modo virtual síncrono mediante la plataforma Zoom®. Además, en cada entrevista se reiteró la solicitud a los participantes para autorizar la grabación de la entrevista.

2.7 Reglas de análisis

Para analizar la operación de sesgos cognitivos durante el proceso de búsqueda de información en internet se elaboraron las reglas de análisis para las consultas, fuentes seleccionadas y resúmenes. En estas reglas de análisis se codificaron los sesgos cognitivos (más comunes) descritos en la literatura, como base para la observación de las trazas del

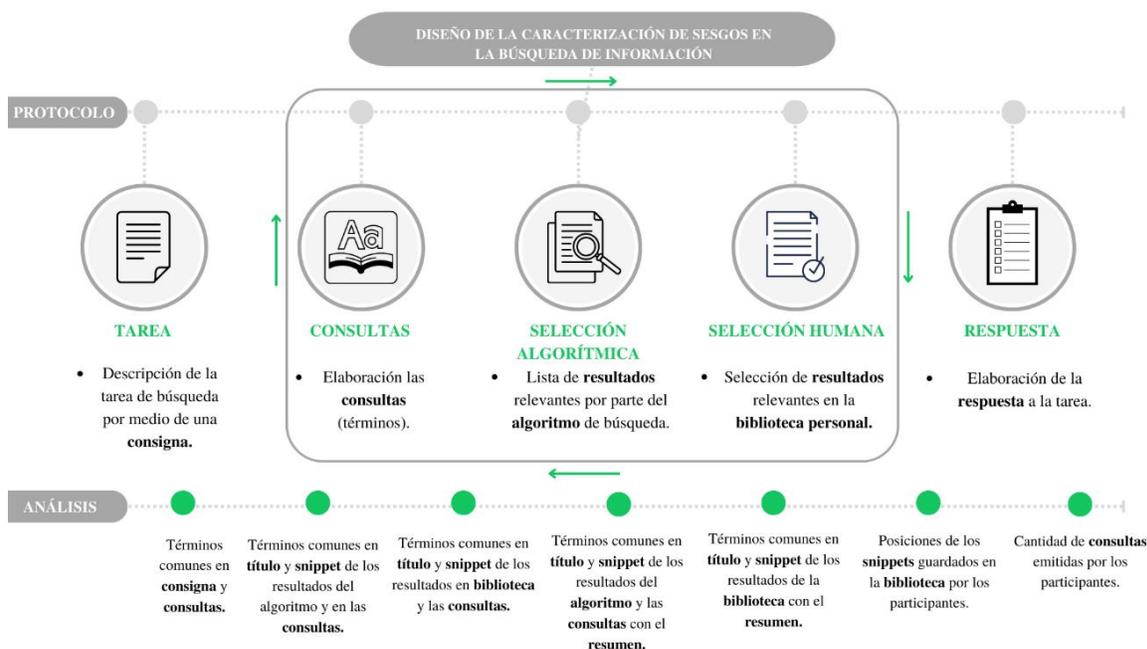
proceso de búsqueda y en las entrevistas, pero teniendo en cuenta la posibilidad de identificar la operación de otros sesgos no considerados previamente.

Tabla 4: Reglas de análisis

CÓDIGO	SCONFIR	SFAMI	SDK	SINFO	SANCLA
DEFINICIÓN BREVE	Tendencia a buscar información que confirme nuestra hipótesis (Matute, 2019).	Sensación de seguridad que predispone a la baja atención. (Matute, 2019).	Sesgo Dunning-kruger es la sobreestimación de las capacidades propias (Kruger, J., & Dunning, D. 1999).	Tendencia a creer que más información implica más veracidad (Baron, 2000).	Uso como punto de referencia del número, valor o información inicial que recibimos (Kahneman, 2012).
DEFINICIÓN COMPLETA	Es la tendencia de las personas a favorecer la información que confirme sus propios presupuestos o hipótesis, sin importar si la información es verdadera.	El sesgo familiaridad con el ambiente o con las personas es definido como una sensación de seguridad que permite bajar el nivel de atención o de alarma (Matute, 2019).	Sobreestimación errónea de su conocimiento o capacidad en un área específica, producto de una falta de metacognición que le dificulta evaluar sus propias habilidades con precisión.	Sesgo de información: Es la tendencia a buscar información, incluso cuando esta no puede afectar a la decisión a tomar. Puede crear la falsa impresión de que, por tener más información, el razonamiento o la conclusión será más veraz.	Sesgo anclaje: La exposición a un cierto número, valor o información fabrica un punto de referencia que influencia la toma de decisiones y la fabricación de juicios relacionados al valor de un objeto. Esto ocurre cuando la percepción de precios de las personas es influenciada por la modificación de los puntos de referencia
CUÁNDO SE USA	Si no hay conocimiento nuevo entre la consulta y la respuesta	Si el alumno siempre utiliza una fuente que conocida (tradicional).	Si un estudiante no utiliza las fuentes de información para la elaboración de su resumen, sino que se basa en su conocimiento y comete errores.	Si un estudiante realiza muchas búsquedas y utiliza muchas fuentes, incluso repetitivas.	Si el estudiante se queda con la información “ancla”, predispone que los resúmenes sean similares a la información inicial.
CUÁNDO NO SE USA	Si hay conocimiento nuevo entre la consulta y la respuesta	Si el alumno utiliza fuentes novedosas.	Si un estudiante utiliza las fuentes de información para la elaboración de su resumen y no comete errores.	Si un estudiante realiza las búsquedas pertinentes y utiliza las fuentes necesarias sin reiteraciones de información.	Si el estudiante no se queda anclado con la información recibida y sus consultas y/o resúmenes no se disponen con la información recibida inicialmente.
EJEMPLO	Busca fuentes de una misma corriente ideológica.	Buscar noticias en fuentes conocidas.	Realiza búsquedas, pero no usa las fuentes y elabora una respuesta errada o parcialmente errada.	Busca gran cantidad de información y la utiliza en su respuesta.	Usa todas o gran parte de las palabras de la consigna de la tarea, asumiendo que son los términos más adecuados para la búsqueda.

Estas reglas de análisis fueron la base para la elaboración del diseño automático de la caracterización de sesgos cognitivos durante el proceso de búsqueda de información (Ilustración 2).

Ilustración 2: Diseño caracterización de sesgos cognitivos en la búsqueda de información



3. Plataforma GoNSA2

GoNSA2 es una plataforma web (Ilustración 3) de búsqueda de información en el contexto de resolución de problemas. El flujo de las funcionalidades de la aplicación se organiza en torno a las tareas o problemas planteados a los estudiantes, lo que brinda contexto a las sesiones de búsqueda. Por lo tanto, brinda a los usuarios una relación semántica entre las consultas, resultados y decisiones. También, permite detener la búsqueda y volver desde el punto en el cual la han dejado previamente.

Ilustración 3: Interfaz de inicio GoNSA2



Para acceder a la plataforma web es necesario que los usuarios se encuentren registrados y cuenten con las credenciales de acceso. Estas credenciales son otorgadas por el administrador de la plataforma y les permiten acceder a la plataforma por medio de un formulario de ingreso, tal como se observa en la Ilustración 4.

Ilustración 4: Interfaz de inicio de sesión GoNSA2

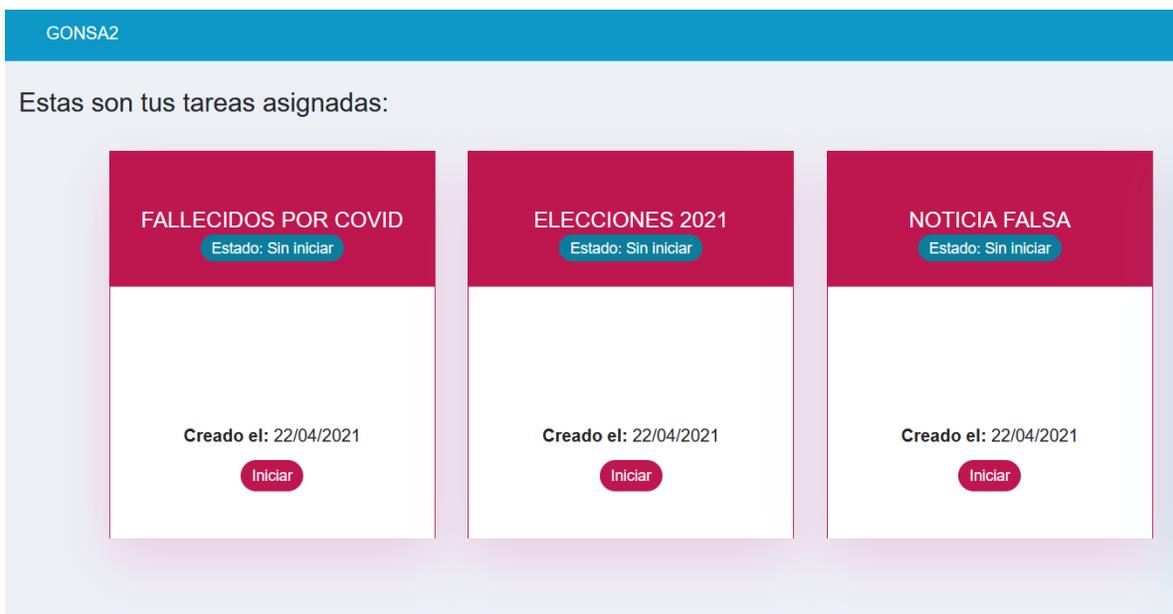
La organización de las tareas planteadas por los investigadores o académicos se encuentran organizadas en categorías para brindar una mayor contextualización y organización a la información presentada en la plataforma. Por ejemplo, en la Ilustración 5 se observa que el usuario tiene solo una categoría asignada, Actualidad.

Ilustración 5: Interfaz de inicio



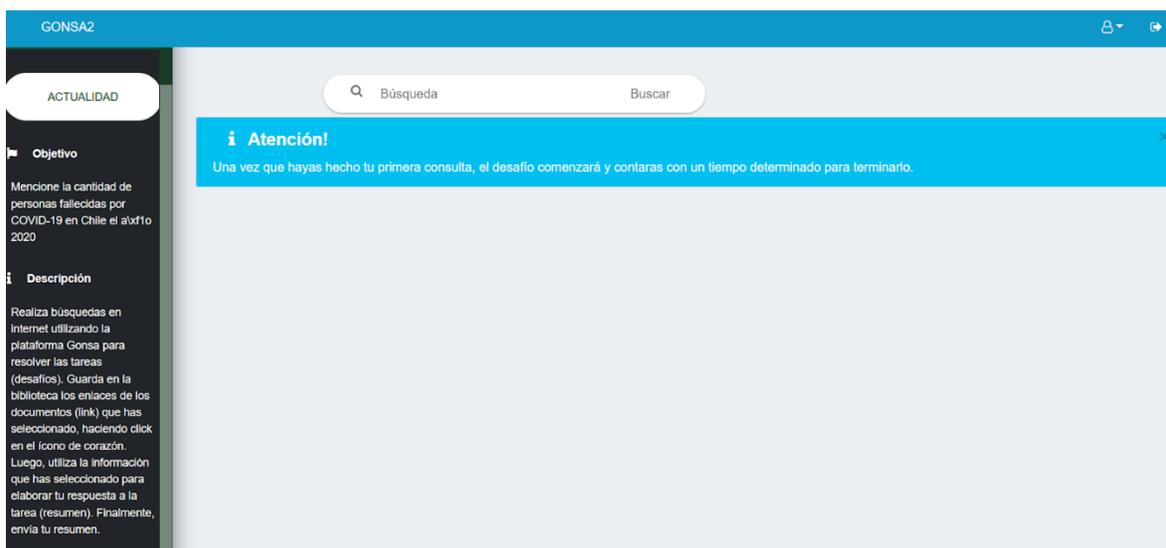
En cada categoría, especificada por los investigadores, se encuentran las tareas que los usuarios deberán resolver. Estas tareas podrán encontrarse en tres estados: “Sin iniciar”, cuando el usuario no ha abierto en ningún momento la tarea; “En desarrollo”, cuando el usuario ha ingresado por primera vez se marca el inicio de la tarea y comienza la contabilización del tiempo de resolución; y “Finalizada”, cuando los usuarios cierran la tarea que han iniciado en algún momento. Estos estados quedan registrados en la plataforma y permiten conocer la línea temporal de acciones de los participantes de los estudios de comportamiento. En la Ilustración 6 se presentan tres tareas que todavía no han sido iniciadas por el usuario activo.

Ilustración 6: Interfaz tareas GoNSA2



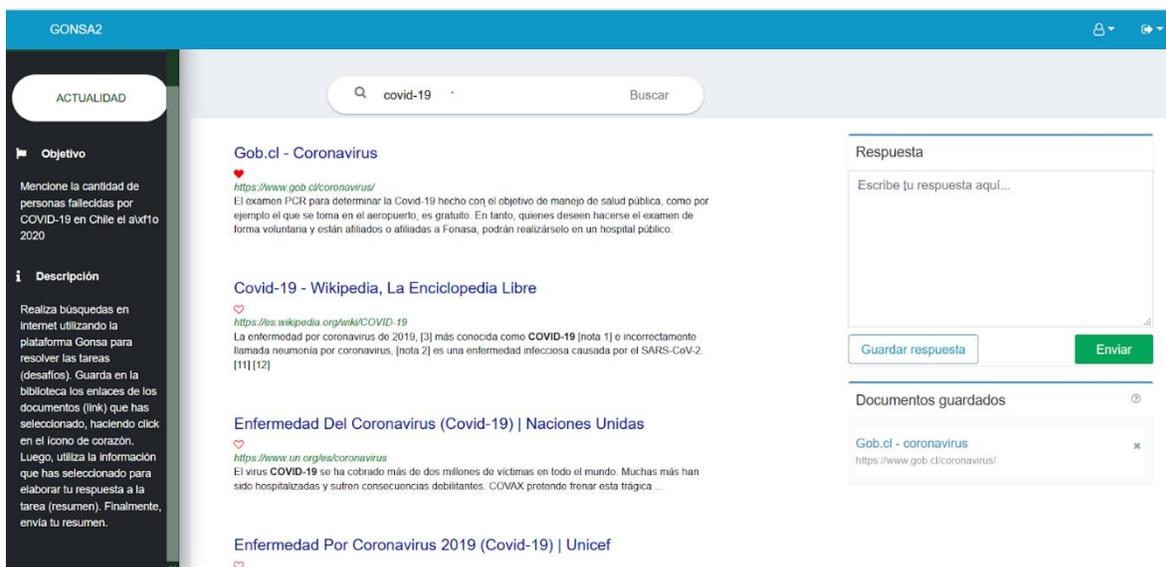
Cuando los usuarios inician una tarea que tienen asignada y que se encuentra activa, estos podrán visualizar una interfaz de búsqueda tal como se observa en la Ilustración 7. En ella podemos ver un menú lateral izquierdo, en el cual aparece la descripción de la tarea que deberá resolver el usuario. En particular, se presenta la categoría, el título, el objetivo de la actividad, una breve descripción y, en algunos casos, una imagen referencial. Asimismo, en la sección central aparece un casillero de búsqueda, el cual permite que los usuarios puedan realizar sus consultas de manera similar al funcionamiento de los principales motores de búsqueda de información. GoNSA2 recibe estas consultas y se conecta con Microsoft Bing para obtener los resultados.

Ilustración 7: Interfaz tarea COVID-19



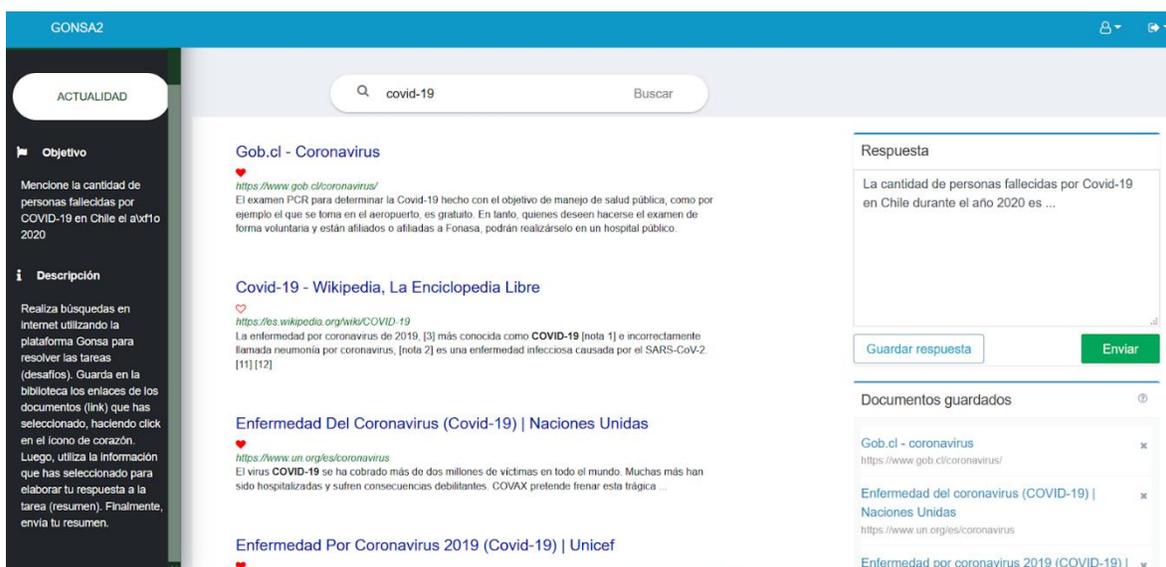
Cuando los usuarios elaboran sus consultas y las emiten en GoNSA2, este recoge los términos para enviarlos a la interfaz de programación de aplicaciones (API) de Bing. Esta conexión permite obtener el listado de resultados de la misma forma que al conectarnos directamente con el motor de búsqueda, salvo que en esta plataforma descartamos la publicidad y le brindamos un contexto a la búsqueda de cada tarea. En la Ilustración 8 podemos observar la lista de resultados obtenida por un usuario activo al enviar la consulta “covid-19” en GoNSA2.

Ilustración 8: Interfaz de búsqueda GoNSA2



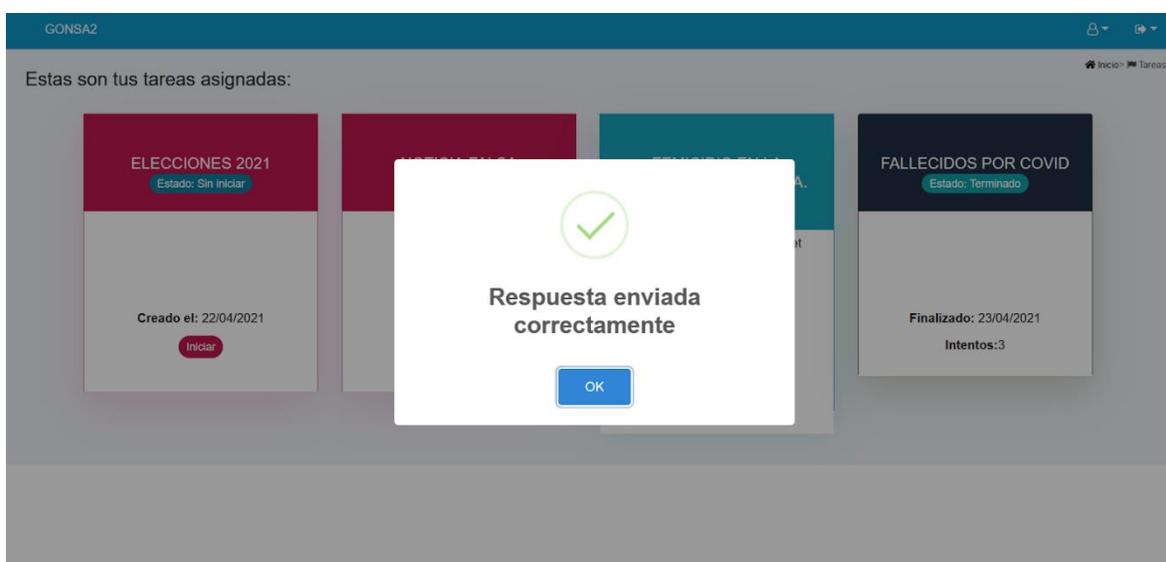
Además, la interfaz activa un casillero de elaboración de la respuesta, la cual puede ir construyendo a medida que avanza en su exploración de la web en base a sus consultas. Adicionalmente, los usuarios pueden marcar documentos como relevantes utilizando el símbolo de un corazón justo debajo de cada resultado, como se aprecia en la Ilustración 9, los que quedan almacenados en su biblioteca personal y pueden ser consultados posteriormente hasta la finalización de la tarea.

Ilustración 9: Interfaz guardar respuesta GoNSA2



Luego de haber explorado la web, emitido consultas, revisado documentos, marcado algunos como relevantes y escrito la respuesta final, el usuario tiene la posibilidad de enviar la respuesta (Ilustración 9) y finalizar la búsqueda. Lo anterior, cierra el proceso de exploración para esa tarea y permite que el participante continúe resolviendo otras tareas dentro de la misma categoría o en otras. En la Ilustración 10 se observa el mensaje de confirmación que se le presenta al usuario cuando termina la tarea, cambiando su estado a “Finalizada”.

Ilustración 10: Interfaz respuesta enviada GoNSA2



GoNSA2 almacena cada una de las acciones de los usuarios de manera anónima, contextualizada y temporal. Es decir, cada decisión del usuario expresada en una acción de búsqueda de información es almacenada de manera anónima con una marca de tiempo que permite modelar temporalmente las acciones, desde el inicio hasta el cierre de esta. Las decisiones que son registradas en la plataforma son las marcas de inicio, las consultas emitidas, los documentos obtenidos de dicha consulta, los documentos que han sido revisados (abiertos) por los usuarios, los documentos que han sido marcados como relevantes, la biblioteca del estudiante para dicha tarea, la respuesta que ha enviado al cierre de la tarea y la marca de cierre de la tarea.

Capítulo Primero

¿QUIÉN ELIGE MI DIETA INFORMATIVA? LA DATIFICACIÓN AUTOMATIZADA DE PERSONAS EN EL AMBIENTE DIGITAL

Vidal-Sepúlveda, M., & Olivares-Rodríguez, C. (2021). ¿Quién elige mi dieta informativa? La datificación automatizada de personas en el ambiente digital. *Información, cultura y sociedad*, (44), 143-154.

doi: <https://doi.org/10.34096/ics.i44.8627>

Indexada en SJR (Q3) y Scopus (Q3)

1. Resumen

Desde la reflexión teórica, este trabajo analiza cómo la tecnología deviene en un dispositivo psicopolítico que configura el nuevo ambiente digital. Primero, como un espacio de flujo de información que otorga sentido a nuestra relación con el mundo y, en una segunda instancia, en la formación de nuestra identidad digital en tanto sujetos en desarrollo que se significan en la experiencia digital. En este contexto, los motores de búsqueda, como software cultural, ofrecen una dieta informativa personalizada y brindan respuestas a consultas por medio de resultados sesgados, elaborados a partir de los patrones de comportamiento de los sujetos y desmitificando, de este modo, el acceso global a la información.

2. Introducción

En la sociedad global el avance de la tecnología informática automatiza la captura de datos sobre las personas, permitiendo la construcción de complejos bancos de datos que reconocen patrones a partir de las acciones realizadas por los sujetos biológicos en la red. Mediante algoritmos se captura, almacena y distribuye su información personal para utilizarla como insumo en el mercado de la información. La datificación algorítmica actúa con mínimas regulaciones, debido a que la opacidad de sus procedimientos dificulta su regulación y fiscalización (Palmiotto, 2021), asimismo, el poder seductor de los servicios globales sobre los sujetos biológicos les permite actuar casi sin oposición de sus usuarios. Esta opacidad de los procesos algorítmicos actúa sobre las personas y los países, afectando la democracia y libertades (Marjanovic & Cecez-Kecmanovic, 2017), en un proceso de datificación que, además, discrimina entre seres y no-seres, es decir, decide quienes son sujetos de interés en la sociedad global. Bajo este esquema existen regiones y poblaciones, dentro y fuera del primer mundo, que no son datificadas, cuyos sujetos biológicos no participan de los “beneficios” ni de las narrativas globales.

En este contexto, los motores de búsqueda emergen como un nuevo mecanismo de control y discriminación del acceso y difusión de la información. Su uso cotidiano oculta el peligro del poder hegemónico que ostentan sobre la población, especialmente la más joven. El aparente libre acceso a la información está determinado por los patrones de comportamiento que los algoritmos como dispositivo psicopolítico han elaborado a partir de nuestras acciones en la red, ofreciéndonos una dieta informativa personalizada. Por lo tanto, no existe un acceso abierto sino uno condicionado por decisiones algorítmicas, aparentemente neutrales e inocuas.

3. La datificación de personas como insumo

La datificación de las personas corresponde a un proceso de caracterización de los sujetos mediante procesos algorítmicos y/o estadísticos como instrumento científico de análisis para la toma de decisiones. Para Mattelart (2007) el proyecto de sociedad de la información toma forma en los siglos XVII-XVIII, en los cuales la matemática representa lo

racional y útil. Este proceso de datificación se iniciaría en la organización del pensamiento a partir del número en la búsqueda por objetivar la sociedad. De acuerdo con nuestro planteamiento, la datificación, como tecnología psicopolítica, describe el comportamiento y los atributos de los individuos y los territorios, siempre que sean de utilidad para la productividad o se encuentren representados en el discurso representativo de los medios (Lodder & Loui, 2018). Debido a esto, grandes masas de población quedan marginadas o invisibilizadas del ambiente digital, bajo una lógica de inclusión/exclusión. Por consiguiente, los sujetos datificados, relevantes para los algoritmos, integran las fuerzas biológicas del ambiente digital y, por tanto, se hace necesaria su caracterización para hacer más eficiente la producción de servicios, lo que les permite controlar no solo a la población, sino también al mercado.

4. La conformación de subjetividades en el ambiente digital

Este nuevo ambiente digital ofrece a los individuos la posibilidad de convertirse en sujetos globales capaces de: conectarse con el mundo, obtener la información que necesitan, compartir y crear nueva información participando de las narrativas transmedia y conformando comunidades (Jenkins et al., 2015). Si bien es cierto que muchas personas pueden estar de acuerdo con esta afirmación, porque en cierta medida se cumplen dichas promesas, esta promesa no es universal ni tampoco gratuita. Este nuevo ambiente es dominado por las empresas globales quienes son dueñas de la mayor parte de los servicios a los que se puede acceder en la web, pero sus transacciones son invisibles para las personas. En este ambiente se transan las acciones registradas por los sujetos biológicos como la materia prima, desde una mirada neoliberal, que es transformada y explotada y, en consecuencia, consumida.

Para la conformación de los que denominaremos “sujetos datificados”, la empresa global utiliza como insumos los datos de los sujetos biológicos que acceden a la red, deseos de participar en un ambiente digital, aparentemente abierto y libre para las distintas formas de comunicación. Por su parte, los países también usan estas interacciones en internet como insumo para el control y vigilancia de sus ciudadanos, ya que la “convergencia digital de todos los datos y tecnologías también amplía al infinito las posibilidades de rastreo y

colonización de las prácticas cotidianas” (Sibilia, 2005, p. 66). De esta forma, el ambiente digital lo domina quién disponga de la tecnología suficiente y apropiada para procesar la big data y elaborar modelos de personalización a partir de las acciones realizadas por los sujetos biológicos en sus redes.

De esta manera, los algoritmos, siguiendo un ordenamiento lógico-racional, consumen las acciones que se desprenden de las interacciones, las que se hacen cada vez más intensas por parte de los sujetos biológicos con los servicios-globales, produciendo sujetos datificados listos para ser consumidos por los algoritmos de personalización, que ordenan la vida, las reglas y la conducta de los sujetos biológicos (Just & Latzer, 2017). Podemos observar cómo servicios-globales de entretenimiento, tales como Netflix (Siles et al., 2019), domestican el comportamiento de los usuarios y los seducen a partir de estrategias de recomendación algorítmicas inferidas a partir de la datificación de las actividades cotidianas capturadas desde cualquier dispositivo. Incluso, los algoritmos y servicios globales replican y potencian los normativos de género presentes en la sociedad (Ni Loideain & Adams, 2020).

En esta inmensidad de datos, solo los algoritmos gobiernan (Sastre Domínguez y Gordo López, 2019), siendo los únicos capaces de extraer patrones y, en consecuencia, guiar el curso de las acciones cotidianas de los seres biológicos, incapaces de procesar datos masivos con diferentes estructuras y definiciones, dando paso a la gobernanza algorítmica (Pasquinelli, 2016). De esta forma, los algoritmos configuran las subjetividades digitales a partir de dichos datos, con el riesgo de amplificar los sesgos de los sujetos y de los algoritmos presentes en la web (Baeza-Yates, 2018, 2020).

En este contexto, las empresas globales seducen a las personas a interactuar con las interfaces de software, quienes están dispuestas a consumir los servicios tecnológicos de generación de información en cada una de sus actividades cotidianas, pero el acceso está condicionado por la brecha digital. Primero, por la brecha socioeconómica que limita el acceso a la conexión a internet a gran parte del territorio mundial y, en un segundo momento, por la brecha de uso, determinada por las capacidades técnicas y de comprensión de los servicios de la web (Llorca-Abad, 2012); una marginación que se acentúa al limitar su posibilidad de participación en las narrativas culturales mediáticas (García-Canclini, 2004), lo que finalmente exacerba la desigualdad social.

5. De la subjetivación del cuerpo a las subjetividades digitales

Tal como lo describe Cachorro “El cuerpo es un mapa, un objeto significativo que adquiere sentidos y significados en sus puestas en escena” (2018, p.3) y, siguiendo este planteamiento, el ambiente digital es un escenario recurrente en el que participan innumerables cuerpos que comparten y nutren sus subjetividades a partir del contenido que producen y consumen. De esta forma, la conformación de subjetividades en el ambiente digital depende de la disposición de los sujetos biológicos a participar activamente en la red, a través de la socialización de su subjetividad contribuyendo a diluir las fronteras entre lo público y lo privado, proyectando su vida personal al ambiente digital (Sibilia, 2012). Para Han (2013, p.68) “los social media y los motores de búsqueda personalizados erigen en la red un absoluto espacio cercano, en el que está eliminado el afuera”. En este nuevo espacio, el sujeto biológico construye una identidad digital que se conforma a partir de sus interacciones en internet, es decir, a los datos que comparte en su red de contactos y los que su red comparte sobre él, por lo que el sujeto biológico pierde el control de su intimidad (Tello-Díaz, 2013), ya que muchas veces no tiene conciencia de los datos que ha registrado, ni tampoco tiene el dominio sobre los datos que se comparten sobre él en sus redes de contacto, puesto que se han desdibujado los límites entre lo privado y lo público, afectando a los cuerpos mediante la digitalización de la vida (Sibilia, 2005). Así, a medida que consumimos los servicios globales, vamos dando forma a nuestra huella digital, la que permanece en la red sin un control sobre quién o quiénes pueden ver nuestros datos ni tampoco los posibles usos de nuestra información (Bazarova y Choi, 2014). Sin embargo, a nuestra identidad biológica le corresponden distintas identidades digitales, construidas a partir de los datos que ha capturado cada uno de los servicios globales que consumimos en la red, complejizando la relación indisoluble entre nuestro cuerpo y nuestra subjetividad (Cachorro, 2008). De esta manera, a partir de un mismo sujeto biológico es posible construir varias versiones datificadas, según el algoritmo y/o el servicio global al que se somete. Por consiguiente, este ambiente digital se constituye de millones de individuos digitales aislados pero conectados. Este homo digitalis descrito por Han (2014a) es un ser expuesto que solicita atención, pero con una identidad disuelta en la masa, incapaz de congregarse ni oponerse al poder, sino tan sólo en movimientos fugaces e inestables. Según su propuesta, no existiría

una clase dominante que explote a este sujeto digital, sino que es el mismo quien se explota y se expone.

Siguiendo la tesis de autoexplotación de Han, la participación activa del ser biológico en la red supone un consentimiento, según el cual el ser biológico vende su individualidad como materia prima a los intereses del mercado de los grandes datos, a cambio de un servicio. Sin embargo, en esta operación de consumo de la oferta global en el ambiente digital no existe una simetría de poder entre los sujetos biológicos y las empresas globales. Es más, planteamos que se accede de manera coercitiva a la construcción de un sujeto datificado por cada servicio global que se consume, ya que muchos de los servicios-globales se ofertan en modalidad gratuita o bajo demanda, relevando como atributos de su negocio la flexibilidad y adaptación a las necesidades e intereses de sus clientes. Además, negarse a la datificación implica la pérdida de servicios globales o bien, la reducción de su calidad y, en consecuencia, ser marginado de las narrativas globales.

Asimismo, a pesar de este supuesto consentimiento, el sujeto biológico desconoce los alcances y las múltiples reconstrucciones a las que es sometido cada vez que realiza acciones en la web y, a su vez, le es imposible caracterizar sus múltiples posibilidades de existencia datificada, desconociendo los trayectos de sus fragmentadas subjetividades (Cachorro, 2008) las que son consumidas por la ciencia de datos para ofertar servicios a la sociedad digitalizada (Igual y Seguí, 2017).

En consecuencia, los algoritmos operan como dispositivos psicopolíticos digitales (Han, 2014b; Vásquez, 2016) que gobiernan por medio de la lectura de patrones utilizados para elaborar predicciones de la conducta de los sujetos biológicos, a partir de sus preferencias en la red. En tiempos de pandemia, el miedo e incertidumbre han facilitado que se incrementen los mecanismos de vigilancia con la complacencia de los sujetos biológicos, a través de dispositivos tecnológicos que capturan datos asociados a la trazabilidad de los enfermos de COVID-19 (Chen et al., 2021; Ho et al., 2021). De este modo, la psicopolítica opera en la web como “técnica de dominación que estabiliza y reproduce el sistema dominante por medio de una programación y control psicológico” (Han 2014b, p.117). Esta sofisticada técnica de poder ya no necesita ejercer la fuerza sobre los cuerpos, sino que el control lo alcanza explotando las emociones humanas a través de la seducción, el análisis y la vigilancia transparente (Han, 2013) de los sujetos datificados.

6. Los motores de búsqueda de información como herramienta para la creación de subjetividades

Actualmente, el acceso a la información se encuentra mediado por una tecnología dominada por las empresas globales y algunos estados hegemónicos - tales como Rusia y China, que cuentan con sus propios motores de búsqueda - colonizando los mercados y la vida cotidiana de todos los sujetos conectados a la red (Li & Dong, 2020). El motor de búsqueda es un software cultural que permite acceder a información indexada en la web y, como tal, corresponde al soporte de acciones culturales, es decir, el motor de búsqueda permite, entre otras acciones:

“crear, intercambiar y acceder a artefactos culturales que contienen representaciones e ideas, creencias y valores estéticos (...); crear y compartir información y conocimiento en línea (...); participar en la ecología de información en línea mediante preferencias y metadatos (...)”. (Manovich, 2013, p.21)

Específicamente, los motores de búsqueda surgen como herramientas creadas para el intercambio de información y conocimiento. Este software cultural ha alcanzado una alta penetración en ámbitos empresariales, educativos y personales y es capaz de ofrecer respuestas a innumerables tipos de consultas, por lo que es una de las aplicaciones de mayor éxito en la red. Estos motores (Google, Bing, Ask, Aol, DuckDuckGo, etc.) pertenecen a empresas globales que poseen índices a la mayor cantidad de información que ha sido posible reunir en la historia de la humanidad (Rasmussen, 2003) y, que los sujetos biológicos disponemos gratuitamente para su distribución a partir de los patrones de selección y exploración de documentos en la web, en una visión utópica del acceso a la información global.

Se suele pensar que al emitir una consulta en un motor de búsqueda es posible acceder a toda la información disponible en la red, sin embargo, consultar a un motor de búsqueda implica la selección de una porción de información de acuerdo con los criterios de la búsqueda, el comportamiento histórico y los mecanismos de indexación (Steiner et al., 2020). Además, no es posible hablar de una red única y abierta, sino que, de acuerdo con su acceso, se la ha clasificado en: internet global, la cual es accesible mediante buscadores y protocolos de uso; internet invisible o profunda, cuya información no tiene indexación automática y es necesario conocer la dirección URL (Fornas, 2003); y la internet oscura, compuesta por redes

restringidas, servidores y redes privadas accesibles mediante autorización o softwares específicos (Gehl, 2014). De acuerdo con esto, el acceso a la información en línea no es global, sino que depende de variables de acceso tecnológico y humano.

Sin embargo, los motores de búsqueda permiten recuperar información disponible en la denominada internet global a través de la vinculación entre los documentos o recursos electrónicos indexados y las consultas emitidas por los sujetos biológicos, según su relevancia (Page et al., 1999), pero la cercanía de los resultados de búsqueda con la consulta emitida depende de variables tecnológicas y humanas, las que al combinarse complejizan el proceso de búsqueda de información. Se debe tener en cuenta que, en el proceso de búsqueda, en la interacción entre la tecnología y el ser biológico intervienen las variables cognitivas, físicas y afectivas de estos últimos (Kuhlthau, 1991), así como a las capacidades de la propia tecnología para responder a las necesidades de búsqueda (Marchionini, 2006).

Dichos algoritmos surgen como una alternativa evolucionista y determinista tecnológica que busca incrementar la eficiencia de las búsquedas frente a la creciente cantidad de información de la nube, disminuyendo el agobio de cientos de resultados con información diversa. La complejidad técnica de los servicios de búsqueda se reduce por medio de algoritmos que investigan los patrones históricos y geolocalizan a los sujetos datificados entregando recomendaciones de consultas (Duarte Torres, et al., 2012), detectando el propósito del usuario (Sadikov et al., 2010; Santos, et al., 2003) o elaborando un ranking que ordena los resultados de la búsqueda considerando criterios tales como la relevancia, la diversidad y la popularidad y el comportamiento histórico del usuario (Balabantaray, 2017; Xing y Ghorbani, 2004), restringiendo y ordenando los resultados de búsqueda a partir de esos filtros (Yoganarasimhan, 2020). Es así, como se ha determinado que los algoritmos seleccionan las respuestas que recibimos de acuerdo con nuestros clics, nuestro idioma y la popularidad del sitio (Jiang, 2014a). En esta ecuación, también suma nuestra ubicación geográfica, denominada por Jiang como “búsqueda provincialista” (Jiang 2014a; 2014b; Cano-Orón, 2019). Asimismo, también está condicionada por la agenda mediática de cada país (Cano-Orón, 2019), la que en países como Chile, es concentrada y controlada por dos grupos económicos. En definitiva, la eficiencia del motor de búsqueda se alcanza mediante la prohibición de acceso a grandes porciones de la web a través de

operaciones heurísticas que descartan caminos de exploración y, con ello, introducen sesgos (Baeza-Yates, 2020) en los resultados.

Para la elaboración de estos modelos, las empresas globales utilizan los patrones de comportamiento de los sujetos datificados con el propósito de dirigir su comportamiento durante la búsqueda y alcanzar mayores niveles de precisión. Los algoritmos hurgan en el banco de datos de sujetos datificados, impactando en las decisiones, en este caso, informativas de los sujetos biológicos. Pero la opacidad de las decisiones algorítmicas (Pasquale, 2015), impiden conocer con exactitud las decisiones del modelo inferido desde el algoritmo, por lo tanto, se ignoran las reglas algorítmicas aplicadas sobre los datos y, en particular, las que seleccionan nuestros resultados en el motor de búsqueda.

Si bien, la empresa global opera en el mundo de los datos a través de algoritmos, los modelos computacionales inferidos desde datos masivos tienen efecto en el mundo biológico de los sujetos a partir de la interpretación subjetiva de los patrones observados, los que modifican su comportamiento (Chander, 2016; Boyd, Levy y Marwick, 2014) y, por tanto, su cultura (Siles et al., 2019), conformando lo que proponemos denominar dieta informativa personalizada. Pero, una dieta informativa dirigida por nuestros patrones de búsqueda no solo restringe el acceso a la información, sino que modela nuestra construcción de subjetividades. Así, la psicopolítica actúa por medio del tráfico de información saturada y filtrada que recibimos diariamente, siendo capaz de generar imaginarios culturales basados en la emoción que reproduce el sistema dominante (Han, 2014b).

En consecuencia, el sesgo de los motores de búsqueda, a través de su ranking de resultados, establece las fronteras de la información, y determina quién tiene o no acceso a cierta información, según esto, estar conectado a internet no es suficiente para asegurar un acceso global a la información. Por lo tanto, la personalización de los servicios-globales de búsqueda no solo reduce el espacio de búsqueda con filtros que determinan los vínculos que, supuestamente, se acercan a nuestros intereses o áreas de estudio, sino que también nos alejan de la diversidad, dejando fuera toda la información antagónica o contradictoria al patrón del sujeto datificado, temporal y geográficamente.

En este espacio virtual, opera también el poder de la empresa global, determinando la información a la que los seres biológicos tienen o no acceso; una lógica de poder seductora que se muestra amigable y accesible, ya que los seres biológicos parecen no sospechar de la

información recibida y mantienen una creencia de que, a partir de la búsqueda, se pueden alcanzar todos los resultados posibles o, supuestamente, los más relevantes. Sin embargo, se ha establecido que los algoritmos de filtro crean una burbuja alrededor de los seres biológicos que limita la oportunidad de acceso a información diversa, su acceso se limita a los sitios que consulta frecuentemente u otros similares (Pariser, 2017). Esto es potenciado por los sesgos cognitivos propios de los sujetos al momento de interactuar con los motores de búsqueda (Baeza-Yates, 2018). Por lo tanto, si el ser datificado nunca o casi nunca explora fuentes de calidad, se reduce su posibilidad de acceder a ellas cuando lo requiera, exacerbando la subjetividad construida a partir de algoritmos y decisiones sesgadas.

Finalmente, la búsqueda de información en la web es una habilidad fundamental en la alfabetización mediática y el activismo de datos que lucha por reducir el control de los algoritmos (Sastre Domínguez y Gordo López, 2019), sin embargo, la mayoría de los seres biológicos no cuenta con las habilidades necesarias para acceder eficientemente a la información (Druin et al., 2009; Qureshi et al., 2015; Şendurur y Yildirim, 2015). Incluso se ha observado una baja tendencia a la utilización de funciones avanzadas en los motores de búsqueda (Yamamoto et al., 2018). De esta forma, el escaso desempeño en esta competencia reduce la eficiencia y efectividad del proceso de búsqueda e impacta en el estado de ánimo de los usuarios (Foo et al., 2014; Rosman, Mayer & Krampen, 2015), dejando muchas decisiones cotidianas a cargo de los algoritmos.

Por lo tanto, la información en la nube no es accesible para toda la población, para su acceso no solo se necesita estar conectado, sino que se requieren habilidades críticas que permitan a los seres biológicos interactuar con la información disponible y, a la vez, habilidades técnicas para estresar al algoritmo y alcanzar hallazgos insospechados.

7. Conclusión

La sociedad global ha modificado las relaciones de poder entre los países y los sujetos, lo que se evidencia en acciones colectivas que han sido capaces de intervenir en la agenda, pero también ha encumbrado a la empresa-global a un nuevo estatus de poder, basado en su capacidad de concentrar y distribuir la información global. Bajo este marco, es urgente considerar los riesgos éticos y legales a los que se ven expuestos los sujetos y los países.

Es necesario reflexionar y tomar acciones frente a la amenaza a la democracia y la reducción de libertades de las personas. El avance de la datificación de personas por medio del uso de servicios-globales, ha permitido concentrar datos personales de los sujetos, en sí mismos sin un valor aparente, pero que son el principio para elaborar sofisticadas bases de datos, a partir la fusión de distintas variables que les permiten obtener información nueva y valiosa para controlar distintos mercados y/o ofertar nuevos productos; es decir, nuestra individualidad, basada en nuestros datos, preferencias, rutinas y deseos, ya no es personal ni de nuestra propiedad. Es más, toda nuestra información disponible en la red es posible de ser: copiada y distribuida por terceros, analizada por algoritmos vigilantes y accesible mediante los motores de búsqueda.

Hoy está en riesgo la libertad y autonomía de las personas para tomar decisiones, desde la selección de un documento en un motor de búsqueda, la compra de bienes y servicios, hasta su influencia en procesos de elecciones. Las personas no tenemos la certeza si nuestra selección está basada en nuestra propia decisión, o si ésta ha sido condicionada por la dieta informativa personalizada elaborada por un algoritmo.

Es por ello que para conquistar la libertad y proteger la democracia en el mundo global, las personas requieren de competencias críticas que les permitan reflexionar sobre sus acciones en el mundo virtual, entendiendo este acceso global como una posibilidad cierta de obtener un mayor acceso a información y de participar democráticamente en las narrativas globales. Una actitud crítica que le permita comprender que el acceso no es global, sino que la globalidad está subordinada a la acción algorítmica que registra, almacena y analiza nuestra información a partir de nuestras acciones en internet.

Por consiguiente, la alfabetización digital es una necesidad y no puede remitirse a la educación formal, sino que debe ser abierta y extenderse a lo largo de la vida. Las personas deben disponer de competencias críticas y técnicas para proteger su individualidad, cuidando su información personal y colectiva, pero, asimismo, apropiarse de los espacios virtuales para elaborar mensajes propios que representen los saberes de las diferentes culturas, actuando en resistencia al poder ejercido por las empresas de las comunicaciones (Castells, 2009). De esta forma, la alfabetización mediática no puede ser reducida al consumo crítico de medios, ya que una participación plena en la cultura digital implica la producción y participación activa de sus miembros. Sin embargo, en nuestra realidad nacional y en el mundo en general, existen

amplios territorios desconectados por su pobreza o aislamiento, cuyos sujetos biológicos no participan de las narrativas digitales. Además, entre los grupos conectados existen deficiencias en las competencias técnicas y de lectura de medios, ambas condiciones necesarias para una participación activa en la cultura de la convergencia mediática.

En consecuencia, la configuración del ambiente digital es desigual y parcial, por lo que la información que puede ser recuperada mediante los motores de búsqueda representa el conocimiento acumulado y la cultura de una pequeña parte de la población mundial. Sumado a esto, para mejorar nuestra experiencia de búsqueda, los modelos algorítmicos seleccionan la porción de datos que se adapta a nuestro perfil, ofreciéndonos una dieta informativa acrítica, eficiente y lista para nuestro consumo.

Capítulo Segundo

WHAT KIND OF SOURCES DO I NEED?: CRITICAL SEARCH FOR INFORMATION ON THE WEB

Vidal-Sepúlveda, M., Olivares-Rodríguez, C., & Cárcamo-Ulloa, L. (2021). What kind of sources do I need? Critical search for information on the Web.

doi: <https://doi.org/10.47989/irpaper913>

Indexada en WOS(Q3), SJR (Q2) y Scopus (Q2)

1. Abstract

Introduction. Previous studies have shown that students have a high confidence in the search engines. This poses a significant risk in learning processes if students do not have critical skills for document selection. This study provides clues about the quality of the information sources that university student's access in their internet searches, and highlights critical thinking as a key competence in personalized information searches.

Method. A quasi-experimental study was conducted with a sample of 58 university students who solved four information tasks. The sources were categorized according to the quality of their editorial process.

Analysis. We stress the critical thinking in a realistic study regarding to both the interactions of university students with not enough validated sources and the type of information task working with.

Results. The main finding of this study is that students mostly access corporate edition sources (52.9%) and non-validated sources (40%). Consequently, the edition type of the source is not related to the ranking elaboration.

Conclusions. The greater access to information does not ensure the quality or veracity of it, confirming the need to promote the development of critical thinking.

2. Introduction

Currently, new technologies have permeated all public and private spaces. The classroom is no stranger to this reality, and therefore, new challenges arise for the education system. Knowledge does not only belong to the teacher, instead it flows and converges in the classroom through different technological devices. The opportunity to access a large amount and variety of information - which grows in unimaginable volumes but, at the same time, does not ensure veracity or quality - complicates the distinction between the useful and the disposable. We move from the certainty of teacher speech and the physical book to the questioning of teacher speech and the uncertainty of online information.

The amount of information offered by the Internet is endless. The indexed web contains at least 5 billion web pages (De Kunder, 2019), many with repetitive and dubious quality content, but with immediate access to hundreds of results (SERP, Search Engine Result Page). This hinders the selection of rigorous information and such overload causes anxiety in students, affecting their decisions (Bawden and Robinson, 2009). Also, the customization algorithms created to facilitate access to information of our interest, adapt SERPs to a profile that search engines create based on our clicks (Pariser, 2017). For example, our Google search is influenced by our Gmail profile information, geographic location and search history.

Therefore, given the increase in information available on technological platforms, it is necessary to have the critical competencies that allow us to perceive it correctly. Today's students need to develop critical thinking skills to: a) integrate new and changing knowledge, b) form their own thinking matrix avoiding the risk of hegemony and homogenization, and c) foster the creation of creative ideas in their area of study and daily life (Olivares-Rodríguez & Guenaga, 2015).

In a personalized search environment - that some people find limiting - it is urgent to characterize the information sources that can be accessed, and thus counteract beliefs about new generations and their mastery of technologies. For example, although new generations demonstrate ease and familiarity with computers, they are dependent on search engine results (Rowlands et al., 2008).

The objective of this study was to determine the quality of the sources of information encountered by students when using a search engine to solve a task, and highlight the need

for critical thinking as a key competence in personalized information searches. To do this, two groups of first-year university students searched for information on the web to solve factual and research / exploratory tasks. To control the bias of previous knowledge, tasks were formulated from different areas and outside the field of study of the participants. Information sources that were accessed through the search engine were categorized according to quality and veracity.

In the first stage of the study, we categorized the information sources according to their content edition in three categories: websites whose content is declared to be edited by corporations or institutions (CORPO), websites with non-validated alternative edited content (ANVAL), and websites with collaborative edited content (COLAB). In the second stage, we analyzed the ranking of the results to compare the different websites in response to the task queries using the RankFlow metric (Rieder, 2016). The study was conducted using the GoNSA2 technology platform (Olivares-Rodríguez, Guenaga & Garaizar, 2018), which implicitly registers all user actions. This quasi-experimental study made it possible to analyze SERPs from a search engine in the same period, comparing the same task developed by students with similar demographic characteristics (age, gender, language, career, and geographical location) but including the singularity of the personal elaboration of the queries.

3. Algorithmic challenges for autonomous learning

The following section summarizes previous findings on critical thinking, bringing together its main definitions and its relationship with student learning. We explain the operation of the search engines and the complexities in the interaction between the student and the personalization algorithms implemented in these.

3.1 Critical thinking as a learning competence

Although there is no consensus on the concept of critical thinking, it is recognized as a superior way of thinking and a necessary competence in the educational field (da Silva & Rodrigues, 2011). Critical thinking has been defined as a complex type of thinking, useful in academia and everyday life. It is also defined as a type of thinking that allows to evaluate

and process in an analytical and reflexive way, creating judgment and accepting or rejecting information of a scientific or daily nature (Tung and Chang, 2009; Yang, 2012). It also allows the creation of new forms of knowledge, while it can also be used to make everyday decisions (Black, 2012; Halpern, 2006).

Psychology highlights the cognitive and self-regulation components of critical thinking, defining it as a superior way of thinking that possesses a set of intellectual skills, aptitudes, and dispositions, the latter of an affective nature (Ennis 1985, 2011). The critical thinker is described as a subject who has developed appropriate standards to evaluate his or her own thinking and who, in turn, uses them constantly to improve their quality and to develop the intellectual characteristics that define the individual as a critical thinker (Elder & Paul, 1994). Scriven and Paul point out that critical thinking is made up of a set of skills to generate and process information, beliefs, and actions based on intellectual commitment (Scriven & Paul, 1992, 2003).

Related to this second approach, Ennis (1985, 2011) defines a concept of critical thinking in which skills are cognitive components and dispositions are affective components, and he or she establishes two dispositions: a) the concern for solidity and b) the concern for justice. In the concern for solidity the critical thinker seeks the truth and that his or her decisions are justified. Faced with the concern for justice, the critical thinker presents his or her position and that of others, showing the foundations and bases of his or her convictions. Ennis (2002) proposes as skills: basic and advanced clarification, fundamentals for decision making, inference, assumption and integration, and auxiliary critical skills, offering a scheme of intellectual operations that are exercised on the information previously collected and selected. Kuhn & Weinstock (2002) also point to metacognitive and epistemological evaluation competencies as essential components in the development of critical thinking. It is no longer enough just to have intellectual competencies and dispositions; the student must also be able to think about his or her own thinking. Daniel (2002) also proposes a critical thinking structure that contains the logical, creative, responsible and metacognitive dimensions.

The Delphi project defined an agreement on the essential skills of critical thinking: analysis, interpretation, self-regulation, inference, explanation, and evaluation (Facione, 1990). According to this, the difference between a critical thinker and a non-critical thinker

is given by these cognitive skills and dispositions. The dispositions are defined by experts as: the attitudes that lead us to use our cognitive abilities to focus on critical thinking. It is worth mentioning that a subject may or may not use his or her critical skills, based on his or her dispositions or motivation (Facione, 2000; Saiz & Rivas, 2008; Valenzuela & Nieto, 2008). This could answer how, when and why people decide to use critical thinking skills, what triggers their use or, conversely, why they choose not to use them. According to this, there are some people with a high disposition to critical thinking and others with a low disposition. The activation of critical thinking would be subject to the personal decision of the subject and not to the absence of the cognitive functions necessary to apply it (Saiz & Rivas 2008; Saiz, Rivas & Olivares, 2015). Therefore, the disposition to critical thinking would determine different ways of facing life, and there would be people who make more thoughtful decisions and others who would not use their full intellectual potential.

A third approach is proposed by The American Library Association, ALA (Berkowitz & Eisenberg, 2003) who analyze critical thinking based on skills related to search, location and use of information. ALA defines nine standards for critical thinking, organized into three groups of basic standards: information literacy, independent learning, and social responsibility. According to this:

- The information literate student enters information efficiently and effectively, evaluates information critically and competently, and uses information accurately and creatively.
- The student, who is an independent learner is information literate and searches information related to his or her personal interests; appreciates literature and other creative expressions of information; and seeks excellence in the search for information and knowledge generation.
- The student, who positively contributes to the learning community and society, is information literate and recognizes the importance of information in a democratic society; practices ethical behavior in relation to information and information technology; and participates effectively in groups to search and generate information.

Therefore, ALA's interest is focused on information literacy as a crucial competence for the subject to access the evidence and construct the arguments that allow the individual to reason logically, away from personal biases that could influence problem-solving. We can see that critical thinking goes beyond scientific thinking, as it must incorporate the affective aspects of the subjects. We conclude that critical thinking is a superior way of thinking, which

requires the direct action of the person who is predisposed to intellectual work. A subject that thinks critically is willing to evaluate, reflect and judge, accept or reject ideas or concepts based on evidence and not on his or her notions or convenience. In addition, he or she constantly monitors the influence of his or her emotions, applying logical criteria and standards, which he or she is constantly willing to question.

3.2 Information search for autonomous learning

Today, access to information is primarily achieved through technology. When students consult a search engine, the information available on the web is retrieved through linked indexed electronic documents or resources, which are displayed according to their relevance (Page, Brin, Motwani & Winograd, 1999). However, the relevance of SERPs in relation to the query made depends on technological and human variables, which together complicate the search for information.

According to the report of the Organization for Economic Cooperation and Development (OECD, 2018), the diversity of pages, the inherent ambiguity and redundancy of the queries, and the low levels of reading comprehension difficult the search process. This reduces the exploration of SERPs and increases frustration, especially in children and young people (Foss & Druin, 2014). It should be noted that in the search process, cognitive, physical and emotional aspects of the students are involved (Kuhlthau, 1991), as well as the ability of technology to respond to the needs of users (Marchionini, 2006).

Models that provide query recommendation (Duarte Torres, Hiemstra & Weber, 2012) and that detect the user's purpose (Sadikov, Madhavan, Wang & Halevy, 2010; Santos, Nguyen & Zhao, 2003) also determine the efficiency of the search process. These, in theory, emerge as an alternative to increase the efficiency of user search against the growing cloud information, reducing the frustration caused by hundreds of SERPs with diverse information. However, search personalization based on user history reduces the search space with filters that determine which results are close to our interests or areas of study, leaving out those that are less related and that present antagonistic or contradictory information.

This 'bubble' (Parisier, 2017) limits access to quality information. A student who explores the web with the purpose of solving a learning task does not have access to all the

diversity of contents (Parisier, 2017) as these will be limited to the sites that he or she frequently visits or to similar ones. Therefore, if the student never or almost never explores quality sources, the possibility of accessing them when required is reduced. For Jiang (2014b), the ranking of SERPs depends on user clicks, but also on language, writing style, site popularity, and geolocation. This last factor is decisive in the SERPs received by the user, as there are significant differences in the information received depending on the geographic location of the user (Jiang 2014a; 2014b; Cano-Orón, 2019).

It is also necessary to highlight the importance of advertising in the search engine business, which coordinates the interests of both content distributors and advertising (Rieder & Sire, 2014). In this sense, the delivery of SERPs through a content ranking is a biased model, because the algorithm determines the priority of some contents over others (Lewandowski, 2017, Rieder & Sire, 2014, Jiang 2014a; 2014b), influencing the access to information. Another factor to consider in search engine biases is media influence, as this also affects the decision making of the search algorithm (Cano-Orón, 2019).

For Haim, Graefe and Brosius (2018), the bubble does not isolate internet users, because personalization does not limit access to all possible options. However, studies that have stressed the bubble have limitations as they are not conducted in real contexts: the information sources are news from a single newspaper, the study considers only a particular type of diversity (Möller et al., 2018), or participants are simulated with profiling algorithms (Haim et al., 2017).

Finally, the search for information on the web is a fundamental skill in media literacy; however, most students do not have the necessary skills to access information efficiently (Druin et al., 2009; Qureshi et al., 2015; Şendurur & Yildirim, 2015).

Research questions

The opportunity to access quality information on the web is conditioned by technology and the critical competencies of the student. If the student makes a greater effort to explore the web and formulate relevant queries, access various sources, select and read these sources, he or she manages to escape the bubble. Therefore, a student who has critical competencies when facing SERPs will select the information to be used according to logical criteria, focusing on the accomplishment of the task, privileging the veracity and quality of the sources. This student will take advantage of the information search technology to achieve

his or her learning objectives, this is why promoting critical thinking is key in hypermediated learning contexts.

The two questions this study seeks to answer are the following:

- Can students access quality information from their queries on the web?
- Do students' critical competencies facilitate access to greater diversity from their web queries?

4. Material and methods

4.1 Participants

The sample was composed of first-year engineering students, between 18 and 19 years old. Informed consent was signed by 58 students out of a total of 65. Of these 58 students, 53 were male and only 5 were female. Data collected were treated confidentially, anonymously and in an aggregated way, and participants were told they could withdraw at any time and without any reprisals.

4.2 Design of information search tasks

Participants received two factual and two research / exploratory tasks (Table 1). Factual tasks refer to specific data search tasks that do not require elaboration by the student but are satisfactorily resolved from sources of authority in the subject. Research / exploratory tasks require greater intellectual effort as the student must look for different sources to solve the task (Marchionini, 2006).

To solve both types of tasks, students had to look for information on the web. They had to evaluate the veracity and quality of the sources reached with their queries. For the development of factual and research / exploration tasks, students had 5 and 25 minutes, respectively. To control bias from previous learning, the topics of the tasks were proposed in different areas of knowledge and outside the formal learning of the participants, considering tasks that have been previously used in the literature (Arguello et al., 2012; Wu et al., 2012; Kules, B. & Shneiderman, B., 2008). In addition, tasks were designed in accordance with the proposal of Wildemuth & Freund (2012).

Tabla 5: Information search tasks

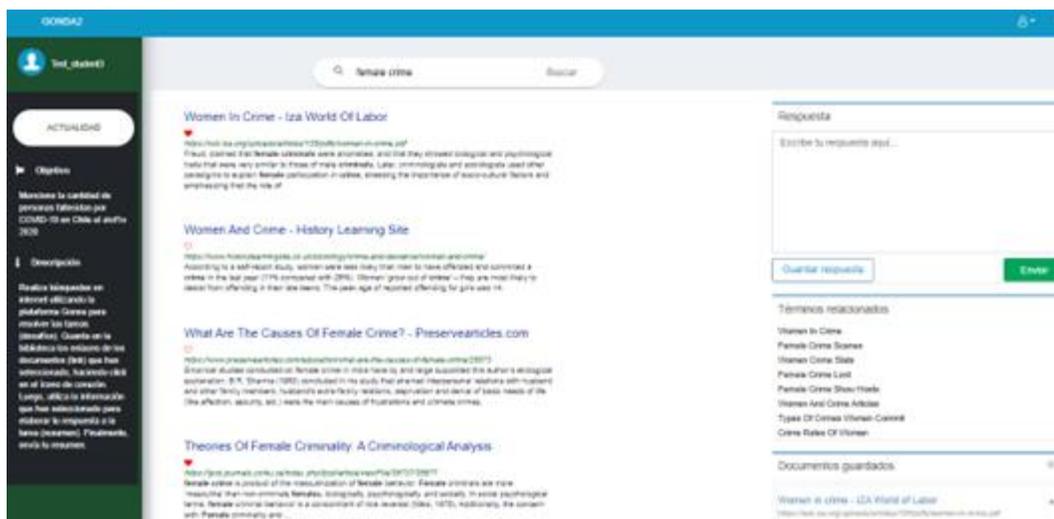
Task	Goal	Type
Sea depth	Find the name of the deepest point in the ocean.	Factual
HIV in Chile	Find the number of people with HIV in Chile.	Factual
Crime in Chile	Determine the components that support a plan to combat crime in Chile.	Exploratory
Build a car	Identify the necessary steps to build a car.	Exploratory

4.3 Materials

Searches were performed using the GoNSA2 technology platform (Olivares-Rodríguez et al., 2018), which interacts with the Microsoft Bing search engine. GoNSA2 is a platform that supports the design of information tasks, the performance of the search, the integration of student behavior and the evaluation of the solutions developed. Illustration 11 shows the information task resolution interface presented to the students, divided in a) the challenge with the goal and description (left), b) the search box and SERP (middle), c) the solution box (right-upper) and d) the personal library (right-lower). The main contribution of GoNSA2 is the generation of detailed information obtained from the search of information, the queries, the documents reached with such queries, and the solutions delivered by the students.

GoNSA2 records all actions performed by users when searching for information to solve tasks. For each task, GoNSA2 records the queries issued by each student, and for each query, it stores the documents and sources delivered by the search engine. It also records the actions that were executed on the results, e.g. if they were seen, stored in the personal library of the task or if they were deleted. In addition, GoNSA2 records the intermediate solutions developed by the students and the time in which the student performed a specific action, from the beginning to the end of the task.

Ilustración 11: GoNSA2, task resolution interface



4.4 Procedure

We presented the functionalities of the technology platform at the beginning of the session. Then, the students had 5 minutes to solve each of the factual tasks, and 25 minutes to perform each of the exploratory tasks. At the end of the session, students had the possibility to solve six optional tasks.

4.5 Source categorization

In order to develop a source categorization, we have considered the edition model behind the websites as a proxy to the information quality assurance. First, we categorized the sources according to the process of content edition in institutional or corporate edition (CORPO), alternative non-validated edition (ANVAL), and collaborative edition (COLAB). CORPO sources correspond to institutional university sites, research centers, specialized press and scientific journals, among others. ANVAL sources correspond to informational sites or sites for the dissemination of miscellaneous content, blogs, etc. that do not have a formal editing process. Finally, COLAB sources correspond to sites such as Wikipedia.

In a second stage, we analyzed the ranking of the results found for each task, to compare the fluctuation of the websites in the queries made, using the RankFlow metric (Rieder, 2016).

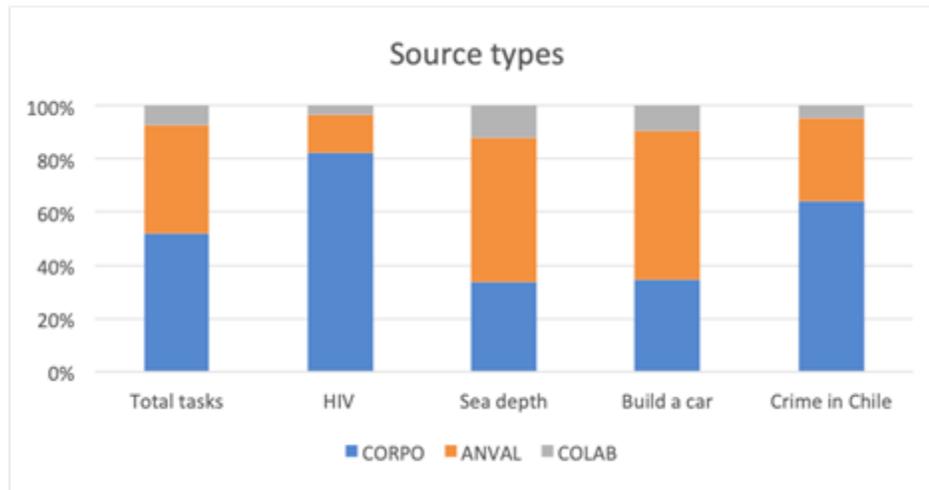
5. Results

5.1 Quality of the information sources according to their process of content edition

To categorize the quality of all the information sources accessed by the students, we reviewed the process of content edition of each of the websites. They were categorized as: CORPO, ANVAL or COLAB. The analysis of source quality was carried out for each of the two participating groups and for each type of task proposed (Illustration 12).

Considering all the proposed tasks, 52.1% of the information sources corresponded to the CORPO category, followed by 40.7% of ANVAL, and 7.2% of COLAB. When we analyzed the quality of the sources according to the type of task requested, the type of source that the students accessed did not seem to depend on the type of task requested, but on the media agenda of the country (Cano-Orón, 2019). For the ‘HIV in Chile’ task, in which the students had to deliver the exact number of people living with HIV in Chile, we obtained 82.1% of corporate sources (Figure 2). This can be explained because this study was carried out in a period of great discussion of the subject, after the media covered the increased HIV contagion and questioned the preventive campaigns in the country. This resulted in increased corporate content (government, specialized press, etc.). Contrary to this, in the factual task ‘depth of the sea’, in which it was requested to find the deepest point of the sea, 54.4% were UNVAL sources (Illustration 12)

Ilustración 12: Source type characterization

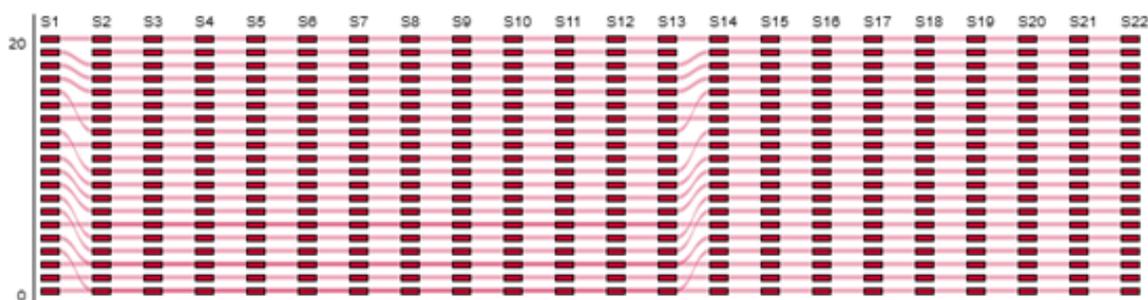


On the other hand, for research / exploratory tasks we raised issues of media interest such as crime and other topics in the scientific / technological field. For the ‘crime in Chile’ task, students had to investigate the web to propose a solution, obtaining 64.1% of CORPO sources, 31.3% of ANVAL sources and 4.6% of COLAB sources. However, for the ‘build a car’ task, mostly ANVAL sources were obtained (56%), followed by 34.6% of CORPO sources (Ilustration 12). It is interesting that COLAB sources obtained a low presence in the SERP rankings for all types of tasks, with only a slight increase in the scientific / technological tasks such as ‘sea depth’ and ‘build a car’. Possibly, the characteristics of the topics do not place them among the most popular searches on the web.

5.2 Website ranking

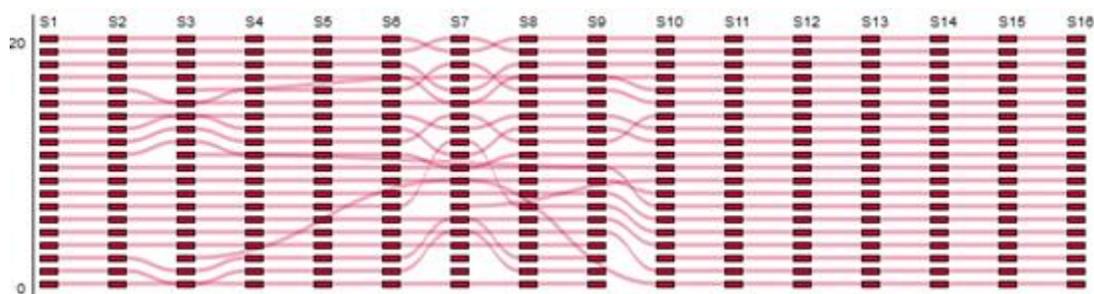
To analyze the fluctuations of the SERP rankings between the different tasks proposed, we compared the rankings using the RankFlow metric (Rieder, 2016), which measures the average deviation distance with a maximum of 1. We considered the first 20 SERPs obtained in each student’s search, for each proposed task. For the ‘HIV in Chile’ task (Ilustration 13), we analyzed 22 students who developed an identical query. The website ranking in Bing showed low fluctuation, as deviations were observed in only two students. The website positioned in the first place was the same for all students. This low fluctuation can be explained because GoNSA2 connects directly and uniquely with Bing. Therefore, the diversity and fluctuation can be only explained by the personal decision of each participant through its search strategies.

Ilustración 13: Variation of the SERPs for the task: 'HIV in Chile'



For the 'build a car' task 16 students searched 'how to build a car'. In this case, the website rankings showed non-relevant fluctuations, which reached 0.09 average distance of deviation from a maximum of 1 (Ilustration 14).

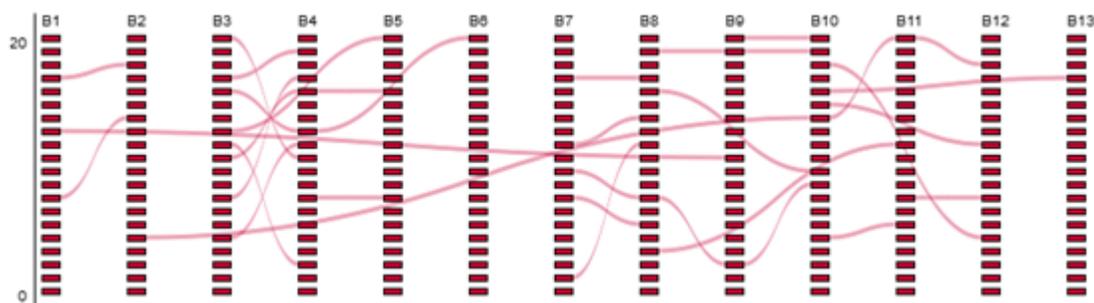
Ilustración 14: Variation of the SERPs for the task: 'How to build a car'



One of the unexpected findings of this study was a student who made 13 different queries to solve the 'Build a car' task, seven in Spanish, five in English and a final one in Spanish. For the same task, the average per student was four queries, privileging the use of the mother tongue with 226 queries in Spanish and only six queries in English. The average deviation distance for the 13 queries of this student was 0.9 (Ilustration 15). But, beyond the fluctuation, the relevance of this finding is that the student elaborated queries that allowed the individual to reach a significant number of new results in each of his or her queries. Also, he or she develop a clearly defined search strategy by querying generic and Spanish key terms and, then, specific and English terms, and finally, back to Spanish but specific terms. Such search strategy provided access to a more diverse set of ordered results against other participants, mainly when English is used (B8-B12). Indeed, he or she obtained different results even in the first place of the web ranking, and he or she reached a greater number of CORPO sources compared to the sample (student: 40.5%, sample: 34.6 %). For this task, the

sample obtained 56% of ANVAL sources. Therefore, search strategy determines the diversity of results, but users are responsible by deciding to move to a deep exploration by questioning the quality of information.

Ilustración 15: SERP variation in the 13 queries performed by a student to solve a research / exploratory task



6. Discussion and outlook

The objective of this research was to determine the opportunity for students to access quality sources when they search on the Internet, considering the biases of the customization algorithms.

We found that most of the websites obtained from the task searches correspond to CORPO sources (52.1%), followed by ANVAL sources (40.7%), and lastly by COLAB sources (7.2%). Although students mostly accessed CORPO sources, there is a high presence of non-validated information, which tests students' critical thinking to select the most appropriate source. It would be interesting to analyse why there is a low presence of COLAB sources in SERPs, considering that the growth of this type of information sources is sustained by the democratization of digital content creation; however, they are not relevant to the algorithm of the search engine. Therefore, user explorations, analysis and decisions mediated by critical dispositions are crucial to improve the relevance of more qualified sources, such as COLAB or authority websites, because the relevance algorithm is based on user actions.

When analysing the quality of the sources versus type of task, the differences observed did not depend on the difficulty of the task (factual or exploratory), but on its topic. We found that the sources reached were mostly of the CORPO type for topics highly covered

by media such as ‘HIV in Chile’ and ‘Crime in Chile’. Coincidentally, both topics correspond to the area of science and technology. As this is a priority area in Chilean education, it is worrying that most of the information that student's access is not of high quality. However, we have categorized the sources based on its editorial process as a quality proxy, but information of high quality and relevance can be also found in authority websites, classified as ANVAL sources. Therefore, user decisions mediated by critical dispositions are crucial to increase the relevance of quality sources by performing deeper explorations of results.

We analyzed the fluctuations of website rankings (RankFlow) elaborated by the search engine Bing, considering only the first 20 websites in order of appearance. For the ‘HIV in Chile’ task, we observed a low fluctuation that in terms of source type showed a greater percentage of CORPO sources, which should ensure quality and veracity in the information. A low variety of sources explains the lower variation in the rankings, which remained practically the same for all participating students. On the other hand, for the ‘build a car’ task, which showed a high deviation in the website ranking, the diversity of the available sources was higher, but this does not ensure the quality or veracity of the contents because most of the websites corresponded to UNVAL sources.

An unexpected finding was a student who elaborated 13 queries to solve the ‘build a car’ task. His or her actions reached different rankings for all the queries, despite their semantic proximity. In addition, the student conducted queries in Spanish and English, which expanded the possibilities of accessing diverse information and increased access to CORPO sources, exceeding the performance of both students’ groups. In the same task, another student made five different queries with only one of them in English, but the low number of queries seemed to prevent significant fluctuations in the website ranking. Same result was observed for the general average of the sample, which only reached 4 queries. This led us to conclude that the number of queries affects the possibility of reaching a greater variety in the ranking of results, because every new query is related to a user perception of not enough information.

We have demonstrated that the type of edition of the source is not relevant for the SERP ranking, unlike other variables such as user history, site popularity, geolocation, language (Jiang, 2014b) and topic media coverage (Suárez-Gonzalo, 2017). Therefore, it is

essential to encourage critical thinking in students to counteract the belief in search engine rankings. This, because students usually accept the first sources that appear in the ranking as valid (Noble, 2018), considering them closer to their query and associating them with quality content and veracity at the snippet level, without clicking the websites.

Personal search strategy defines the diversity and quality of sources which are provided by the algorithms of search engines. Such strategy is mainly based on personal decisions of which terms are used, what sources are explored, how many queries are submitted and which languages are used. Such interactive decisions are considered by the relevance algorithms of search engine by providing more diverse results. Therefore, users with critical dispositions during sources analysis provide a strong relevance feedback in order to support quality sources non explored previously.

Considering that other studies confirm the use of search engines by young people to solve their academic tasks and perform leisure activities (García-Martín & Cantón-Mayo, 2019; Rowlands et al., 2008), it is key to take actions to improve their critical thinking skills. This will allow students to overcome the personalization of search engines and access quality content. It seems particularly necessary to educate about information search in the first years of university.

We have defined some pedagogical guidelines: a) define a first year course at degree level about information search with a critical view, which is focused on media and information literacy, source credibility and algorithmic decisions of search engines, b) promote to students a critical analysis about personalization and recommendation algorithms behind search engines, c) promote to students a properly use of quality sources since first year and d) promote a search with an inter-lingua strategy to reach a higher diversity results during english and spanish courses.

To continue with this study, we will expand the sample to university students from other careers and other countries, and will compare the results obtained with other areas of knowledge and with the geographical location of the queries. We also plan to conduct a new experiment in another search engine such as Google, to study the possible differences in the students' search results, its influence in the web browsing behavior and in document selection

criteria and solution elaboration. Finally, we are planning a longitudinal study in order to analyze the search strategy evolution along the different levels at a degree program.

Capítulo Tercero

BUSCAR Y SELECCIONAR INFORMACIÓN EN INTERNET: SESGOS, SATURACIÓN DE INFORMACIÓN Y PENSAMIENTO CRÍTICO

1. Resumen

El crecimiento exponencial de la información en internet y la indexación algorítmica han complejizado la búsqueda de información y el razonamiento crítico. El presente estudio busca caracterizar los sesgos cognitivos y el comportamiento de los estudiantes universitarios durante la búsqueda de información en internet. Se realizó un estudio experimental con 64 estudiantes (1er año universitario) que resolvieron tres tareas de búsqueda de información. La plataforma GoNSA2, recopiló automáticamente las trazas del proceso de búsqueda para cada tarea. Los principales hallazgos refieren a la operación de los sesgos cognitivos de anclaje, ranking y disponibilidad, los que dependen de la operación del pensamiento crítico, la saturación de información y del contexto de las tareas. Los estudiantes tienden a anclarse con la consigna de la tarea para seleccionar las palabras de sus consultas: 9 de cada 10 estudiantes para la tarea “Elecciones 2021” y 5 de cada 10 para la tarea “COVID-19” y “Noticias falsas”. También, se evidencia que los estudiantes seleccionan las fuentes de información basándose principalmente en el título del resultado que ofrece el buscador. Estos hallazgos nos permiten sustentar la importancia de modelos de búsqueda que habiliten nuevas estrategias pedagógicas orientadas a fomentar la competencia crítica de búsqueda de información.

2. Introducción

Para solucionar una necesidad de información accedemos a internet a través de nuestro motor de búsqueda favorito, desde cualquier lugar con conexión y desde diferentes dispositivos tecnológicos. A este aumento en el acceso a la red se suma el crecimiento exponencial de la información en internet y la indexación algorítmica, complejizando el razonamiento crítico y la recuperación efectiva de información fiable. La búsqueda de información en internet está sujeta a las reglas de la web. De este modo, recuperar información requiere de la mediación de un algoritmo de búsqueda. En internet existen diferentes motores de búsqueda y cada uno de ellos cuenta con una mediación algorítmica que tiene sus propios criterios de indexación y recuperación de información. De esta forma, no ha sido posible determinar si, ante una consulta idéntica, es probable obtener la misma fuente de información o el mismo orden de relevancia en diferentes motores de búsqueda (Yagci, et al., 2022). Actualmente, estos algoritmos son capaces de inducir modelos generalizados de comportamientos o fenómenos, a partir de grandes cantidades de datos (Kroll, 2015; Caplan y Boyd, 2016; Baeza-Yates, 2018). Existe una corriente de investigación preocupada por las implicancias de los modelos computacionales inferidos desde datos masivos sobre la vida cotidiana de las personas, debido a las modificaciones del comportamiento (Chander, 2016; Boyd et al., 2014). Lo anterior, ha incrementado su tratamiento en la literatura, traspasando las fronteras de la informática para ser un objeto de interés de las ciencias sociales, principalmente en lo relativo al uso, manipulación y privacidad de la información (Burrell, 2016, p.2). Dichas decisiones algorítmicas han sido definidas como opacas (Pasquale, 2015), debido a que no es posible conocer con exactitud las decisiones del modelo inferido desde los distintos algoritmos, por lo que las personas ignoramos las reglas algorítmicas aplicadas sobre los datos y, en particular, las que seleccionan nuestros resultados en el motor de búsqueda. En consecuencia, la mediación algorítmica de estos dispositivos tecnológicos acota los resultados de manera transversal a todos los usuarios, limitando el acceso a la información disponible en función de sus decisiones históricas, la geolocalización y el idioma entre otros criterios (Jiang, 2014; Pariser, 2017; Olivares-Rodríguez et al., 2021; Vidal-Sepúlveda et al., 2021).

En este ambiente digital los estudiantes acceden a la información a través de la selección del algoritmo del sistema de búsqueda, devenido en curador mecánico de

contenidos (Bhaskar, 2018). Esto supone la oportunidad de acceder a fuentes de conocimiento de difícil acceso antes de la digitalización. Podríamos asumir que el vasto uso de la tecnología como medio de comunicación faculta a los estudiantes como expertos en la búsqueda de información. Sin embargo, tienden a confiar en la tecnología, delegando sus decisiones en las recomendaciones del algoritmo de su buscador favorito, por ejemplo, aceptando como válido el resultado que más se repite (Ghenai et al., 2020).

Durante la pandemia, aumentó la difusión de información sesgada, con errores e incluso falsa (Gutiérrez-Coba et al., 2020). Asimismo, la educación avanzó forzosamente a modelos de enseñanza virtual e híbrida, lo que aceleró la irrupción de las TIC en las instituciones educativas, dejando manifiesta las brechas digitales de los docentes y estudiantes. En general, para los estudiantes, comunicarse y aprender de manera ubicua es una actividad integrada en la vida cotidiana (Cope y Kalantzis, 2010; Burbules, 2012). En este sentido, y tal como lo plantea Pereira et al. (2019) los estudiantes desde edades cada vez más tempranas participan activamente del entramado mediático global. No obstante, en su estudio evidencia que los medios aún no están integrados en el aula. Por lo tanto, los estudiantes quedan expuestos a la sobrecarga de información, la desinformación, las noticias falsas, entre otros desórdenes informativos. Esto supone la necesidad urgente de formar a los estudiantes para la selección cuidadosa de la información. Sin embargo, desde las instituciones educativas no existen planes de formación de alfabetización crítica que les permitan seleccionar fuentes fiables de información ni cuestionar el contenido de la información a la que acceden en la web. Por ejemplo, en un reciente estudio de Sales et al. (2020) se evidencia que el profesorado universitario español manifiesta preocupación respecto a las competencias críticas de sus estudiantes. Asimismo, cuestionan sus propias capacidades para formar a sus estudiantes en estas competencias. En general la alfabetización digital y mediática de los jóvenes queda a su propia agencia, aprendiendo, principalmente, a partir del ensayo y error, imitación e inspiración y la búsqueda de información (Pereira et al., 2019). No obstante, en la práctica no se ha avanzado lo suficiente hacia la formación de docentes que incorporen la formación de una perspectiva crítica en su práctica pedagógica (González-Osorio, 2021; Mesquita-Romero et al., 2022) que provea a las personas de las competencias críticas para desenvolverse como ciudadanos autónomos capaces de realizar

una comprensión crítica de los medios y la información disponible (Buckingham, 2008; Gozávez-Pérez y Aguaded-Gómez, 2012).

Antecedentes

De acuerdo con la teoría de racionalidad limitada de Simon (1957) el proceso mental humano es limitado “(...) las personas eligen y razonan racionalmente, pero solamente dentro de las restricciones impuestas por su búsqueda limitada y sus capacidades de cálculo.” (Cortada, 2008, p. 69). A partir de esta teoría, Tversky y Kahneman (1983) plantean su teoría de pensamiento dual, según la cual las personas contamos dos formas de pensamiento: el sistema 1 y el sistema 2. El sistema 1 es el más antiguo y toma de decisiones basado en la intuición. Cortada (2008) explica que este sistema se activa de forma automática y se conduce mediante el hábito. Este automatismo le permite tomar decisiones en ambientes de incertidumbre y ser eficiente en multitareas, sin embargo, dificulta su control. Por el contrario, el sistema 2 corresponde a lo que conocemos como pensamiento racional y es el más reciente en la evolución humana. Se caracteriza por ser lento, controlado y por el gran esfuerzo cognitivo que exige. Este tipo de pensamiento requiere ser activado, lo que se consigue mediante el entrenamiento constante y prolongado (Matute, 2019). Lo anterior, facilita que las personas utilicemos los juicios intuitivos la mayor parte del tiempo, sobre todo, en ambientes de incertidumbre. Según Matute (2019) el problema surge cuando estos atajos cognitivos nos conducen a sesgos cognitivos, debido a la urgencia de elaborar juicios inmediatos para diferentes estímulos o situaciones y, a la imposibilidad de procesar rápidamente toda la información disponible. En concreto, los sesgos cognitivos corresponden a formas no razonadas para tomar decisiones con bajo nivel de esfuerzo cognitivo, y que pueden desencadenar la toma de decisiones parciales o erradas (Tversky y Kahneman, 1983), que podrían afectarnos a nivel individual y colectivo.

En la literatura es posible encontrar diversos autores que exponen la ocurrencia de sesgos cognitivos en internet (Kroll, 2015; Caplan y Boyd, 2016; Baeza-Yates, 2018). Baeza-Yates (2018), describe una serie de sesgos en la búsqueda de información, tales como la indexación, el muestreo, el ranking, la interacción, la selección y la interacción social. Estos sesgos derivarían de una dinámica sistémica compuesta por las decisiones cognitivas de las

personas en su interacción con la tecnología y la sociedad, a lo que se suman las decisiones algorítmicas de personalización y selección de la información. Durante una sesión de búsqueda educativa los sesgos humanos ocurren en tres momentos clave: a) La elaboración y/o reformulación de una consulta; b) La decisión de seleccionar y/o guardar una fuente de información y; c) La construcción de una respuesta a la tarea. Asimismo, los sesgos algorítmicos surgen a partir de acciones humanas. Primero, por la intervención del programador (Baeza-Yates, 2018) y, segundo, por: d) la ejecución de sesgo algorítmico para cada consulta de usuario. Por lo tanto, el sesgo algorítmico es alimentado por las búsquedas y se materializa en una lista de resultados conformada por un título y un *snippet* (resumen del resultado que incluye el título y una breve descripción).

Se encontraron estudios recientes que analizan los sesgos en el comportamiento de los usuarios, a partir de la manipulación de los resultados de búsqueda (Novin y Meyers, 2017), las decisiones de los usuarios (Ghenai et al., 2020), el tiempo invertido en la búsqueda (Pothirattanachaikul, et al., 2019) y en la evaluación de los resultados (Liu, 2019). Estos estudios confirman que a partir de la caracterización del comportamiento de búsqueda de las personas es posible identificar la operación de los sesgos cognitivos inducidos por los motores de búsqueda. En cuanto a la influencia de la saturación en el comportamiento de búsqueda, se ha determinado que los estudiantes que reciben resultados no saturados visitan más páginas y dedican más tiempo a la exploración de la web que quienes recibieron resultados saturados (Kelly y Azzopardi, 2015). Sin embargo, no se encontraron estudios que establezcan la relación de la saturación de información sobre la operación de los sesgos cognitivos derivados de las decisiones de los usuarios en contextos naturalistas.

En esta investigación buscamos indagar en el uso del motor de búsqueda como la principal plataforma de búsqueda de información en internet. Entendiendo que, a pesar de su difundido uso, las personas no conocemos los criterios de indexación de los algoritmos de los buscadores ni cuestionamos el alcance de la influencia del sistema de búsqueda en lo que denominamos dieta informativa personalizada. Asimismo, entendemos que optimizar las ventajas de la búsqueda de información en internet resulta un proceso clave en el aprendizaje autónomo en todos los ciclos y tipos de formación. Por lo anterior, nos interesa recopilar evidencias de los obstáculos que enfrentan estudiantes universitarios durante la búsqueda de información para resolver tareas académicas. Nuestro objetivo es establecer la relación causal

que tiene la interacción de la saturación de información en los resultados de búsqueda y la motivación al pensamiento crítico, sobre la operación de los sesgos cognitivos durante la búsqueda de información en internet. Por ello, analizamos el comportamiento de búsqueda de los estudiantes, caracterizando los posibles sesgos y los efectos de la saturación de información, considerando la operación de la motivación de los estudiantes a pensar críticamente. En ese sentido, esta investigación busca responder las siguientes interrogantes:

- ¿Cómo operan los sesgos cognitivos durante la búsqueda de información para resolver una tarea académica?
- ¿La motivación de los estudiantes a pensar críticamente influye significativamente sobre la operación de sesgos cognitivos?
- ¿Los estudiantes que reciben una lista de resultados saturada manifiestan un comportamiento de búsqueda más sesgado?

Estas preguntas serán abordadas mediante los datos proporcionados en la sesión experimental con estudiantes universitarios para determinar los posibles sesgos y la influencia de la saturación y la motivación de los estudiantes a pensar críticamente.

Material y Métodos

3. Material y métodos

3.1 Participantes

El universo de estudio corresponde a estudiantes universitarios. Se utilizó una muestra no probabilística por conveniencia (Flick, 2007). El criterio para seleccionar las unidades de análisis es el nivel de formación educativa de los sujetos, por lo que se selecciona a los estudiantes de primer año, ya que tienen un avance similar en la malla curricular de su carrera. La selección de la muestra corresponde a un proceso aleatorio de la población. La muestra corresponde a 64 estudiantes universitarios de primer año de las carreras de periodismo e ingeniería en informática. La media de edad es de 21,6 años, con un 51,6 % de hombres y 48,4 % de mujeres. Los estudiantes comprometieron su participación de manera voluntaria mediante un consentimiento informado en el que autorizan el tratamiento de sus datos de manera confidencial, anónima y agregada con fines estrictamente académicos. La

aleatorización de la muestra se realiza en forma automática en la plataforma GoNSA2 (Olivares-Rodríguez et al., 2017), la que interactúa con el motor de búsqueda Bing de Microsoft, registrando las acciones realizadas por los estudiantes durante la búsqueda de información.

3.2 Diseño y procedimiento

Esta investigación se enmarca en un estudio experimental (Campbell y Stanley, 1995). Se conformó un grupo experimental que recibió una lista de resultados saturada (20 ítems por página), mientras que el grupo de control recibió una lista de resultados tradicional (10 ítems por página). La asignación de participantes a cada grupo se realizó de manera aleatoria (Campbell y Stanley, 1995; Greenwood, 1973) siguiendo un proceso estratificado aleatorio en función de las variables de 1) motivación al pensamiento crítico y 2) área disciplinar, lo que permite distribuir la muestra de manera homogénea en los grupos. Ambos grupos realizaron un test ad-hoc para medir la motivación al pensamiento crítico (Valenzuela y Nieto, 2008). El diseño experimental considera a la personalización algorítmica y el área disciplinar como variables extrañas por lo que fueron controladas durante el estudio. La personalización algorítmica se controló mediante un punto único de conexión a través de la API de Bing y, el área disciplinar, considerando a estudiantes de primer año de dos áreas disciplinares. Asimismo, la aleatorización de ambos grupos permite igualar otras posibles variables extrañas desconocidas por los investigadores (Pick y López, 1998).

3.3 Variables independientes

Motivación al Pensamiento crítico: Se utilizó la Escala Motivacional de pensamiento Crítico, en adelante EMPC (Valenzuela y Nieto, 2008). La EMPC se compone de 20 ítems tipo Likert con puntuaciones de 1 a 6, en la cual los sujetos expresan su grado de acuerdo o desacuerdo de una serie de afirmaciones respecto de las expectativas que tienen de pensar críticamente y, de afirmaciones valorativas respecto la importancia y utilidad percibida hacia el pensamiento crítico, el costo que asumirían por hacerlo y su interés personal en pensar de forma rigurosa. Así, los ítems de la EMPC son: Expectativa, Valor, Utilidad, Interés y Costo, los que se distinguen factorialmente en 5 subescalas en 1er y 2do orden ($KMO= ,887$ y $,902$),

correspondientes a los constructos propuestos por el modelo de Eccles y Wigfield (2002) con altos niveles de confiabilidad. Para el ítem Expectativa ($\alpha=,774$), Importancia ($\alpha=, 770$), Costo ($\alpha=, 775$), Utilidad ($\alpha=, 790$) e Interés ($\alpha=, 724$).

Saturación de Información: Corresponde a la sobrecarga de información que reduce la capacidad de tomar decisiones efectivas por parte de los usuarios (Faulkner y Trotter 2017; Roetzel, 2019). En particular, durante la búsqueda de información esta sobrecarga se evidencia en la cantidad de resultados por página o en la cantidad total de resultados (Kelly y Azzopardi, 2015).

3.4 Variables dependientes

Operación de sesgos cognitivos: Para la identificación de la operación de los sesgos cognitivos (Tversky y Kahneman,1983) se propone observar su operación en los diferentes puntos de toma de decisiones por parte de los estudiantes en el proceso de búsqueda: 1) las consultas emitidas por los estudiantes, 2) el contenido de los documentos seleccionados en cada una de las sesiones de búsqueda de información en la web, 3) el contenido de los resúmenes elaborados. De esta manera, la caracterización de cada uno de los sesgos se detalla en el apartado correspondiente.

3.5 Diseño de tareas de búsqueda

Para seleccionar las tareas de búsqueda se elaboró un conjunto de actividades siguiendo los lineamientos de la taxonomía bidimensional de Anderson y Krathwohl (Urigo et al., 2019). Las tareas fueron sometidas a la evaluación de tres académicos expertos en: Educación y Tecnología, Neurodidáctica y Lingüística. Se evaluó la comprensión de la consigna, el tipo de conocimiento, el proceso cognitivo, la mediatización, las fuentes de información disponibles y el tiempo de ejecución. En el experimento se utilizaron las tareas que alcanzaron acuerdo en todas las dimensiones analizadas por los expertos. En dichas actividades se invita a los estudiantes a elaborar una respuesta fundamentada mediante una búsqueda de información en la web (Tabla 6).

Tabla 6: Tareas de búsqueda de información

Contexto	Consigna de la tarea	Proceso cognitivo	Tipo de conocimiento
Fallecidos COVID 19	Mencione la cantidad de personas fallecidas por el COVID 19 en Chile durante el año 2020.	Recordar	Factual
Elecciones 2021	Resuma las razones del cambio de fecha de las elecciones de abril del 2021.	Comprender	Exploratorio
Noticias Falsas	Resuma las características de las noticias falsas.	Comprender	Exploratorio

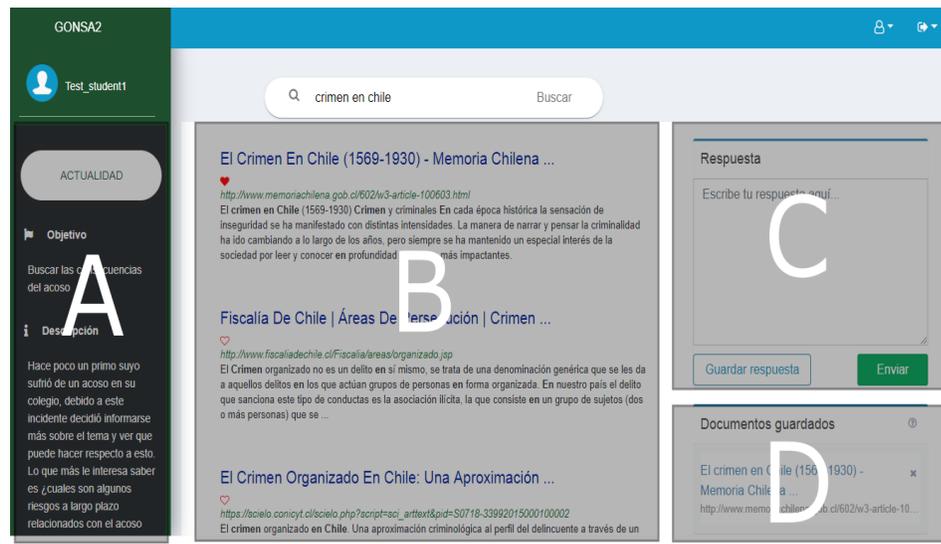
Fuente: Elaboración propia

3.6 Técnicas de recolección de datos

La comunicación con los estudiantes se realizó de forma síncrona-online (entre el 22 y 26 de abril de 2021). En la primera sesión se aplicó la escala EMPC (Valenzuela y Nieto, 2008). Para determinar el nivel de motivación al pensamiento crítico de los participantes se consideraron las medias de las respuestas en cada dimensión del EMPC (Valenzuela y Nieto, 2008) por cada estudiante.

En la segunda sesión se presentaron las funcionalidades técnicas de la plataforma y se les invitó a resolver las tareas utilizando el buscador de la plataforma (B) en el orden de aparición siguiendo las instrucciones de cada tarea (A) y guardando las páginas webs que consideran relevantes en su biblioteca personal (D). Finalmente, debían enviar su respuesta (C) a la plataforma (Ilustración 16). Los estudiantes dispusieron de 45 minutos para resolver las tareas.

Ilustración 16: GoNSA2: Interfaz de resolución de tareas de información



Fuente: Elaboración propia

3.7 Caracterización de sesgos cognitivos

En esta investigación caracterizamos los posibles sesgos durante una sesión de búsqueda de información usando una estrategia de análisis ordenada a partir de los momentos clave según la secuencia de realización de las tareas de búsqueda (Ilustración 17).

Ilustración 17: Diseño de la metodología de caracterización de sesgos cognitivos



Fuente: Elaboración propia

4. Resultados

4.1. Sesgo de anclaje

Según Kahneman (2012), el sesgo de anclaje corresponde al uso como punto de referencia del número, valor o información inicial que recibimos. Para determinar el sesgo de anclaje de los estudiantes en las tareas de búsqueda analizamos la coincidencia de palabras en las relaciones previamente establecidas en el diseño de análisis (Ilustración 2). Consideramos que en los estudiantes opera el sesgo de anclaje si existen coincidencias de palabras o términos o, por el contrario, utilizan palabras o términos relacionados. De esta forma, el sesgo de anclaje (SA , ecuación 1) es la proporción entre el número de palabras contenidas en cada relación (P_c) y el número total de palabras (P_q). Para mejorar la precisión del análisis se eliminaron las tildes y *stopwords* que no pertenecen al léxico nocional (conectores, pronombres, artículos, etcétera).

$$SA = \frac{P_c}{P_q}$$

Durante la sesión de búsqueda, los estudiantes tenían la posibilidad de seleccionar resultados que, a su juicio, fueran relevantes para alcanzar el objetivo de cada tarea, almacenándose en su biblioteca personal (Ilustración 16). Para caracterizar el sesgo de anclaje fueron considerados sólo los *snippets* de los resultados seleccionados, ya que son la porción de información que influye directamente en la decisión temprana de los participantes tanto para explorar como para seleccionar información. Por lo tanto, la relación de cada uno de los *snippets* guardados en la biblioteca personal con el resumen elaborado por los estudiantes nos brinda información acerca de cuánto de esas páginas fue considerado relevante por los participantes y formó parte de sus respuestas.

4.1.1 Caracterización del sesgo de anclaje

Elaboración y/o reformulación de las consultas:

En esta etapa observamos la operación del sesgo de anclaje en la relación “consigna - consultas” en cada tarea (Tabla 7). En general, se observa que el sesgo de anclaje es mayor en los estudiantes con mayores resultados en la escala EMPC.

Tabla 7: Sesgo de anclaje entre la consigna - consultas

Dimensión del EMPC		Covid-19		Elecciones 2021		Noticias falsas	
		Experim.	Control	Experim.	Control	Experim.	Control
Expectativa	Alto	0,47	0,59	0,92	0,91	0,44	0,58
	Bajo	0,41	0,51	0,74	0,74	0,42	0,51
Importancia	Alto	0,43	0,54	0,76	0,75	0,53	0,45
	Bajo	0,39	0,4	0,81	0,77	0,27	0,78
Utilidad	Alto	0,42	0,56	0,81	0,76	0,38	0,56
	Bajo	0,44	0,39	0,62	0,77	0,58	0,37
Interés	Alto	0,51	0,47	0,78	0,69	0,46	0,58
	Bajo	0,36	0,59	0,79	0,83	0,39	0,42
Costo	Alto	0,5	0,51	0,68	0,72	0,33	0,63
	Bajo	0,41	0,53	0,8	0,78	0,44	0,46

Fuente: Elaboración propia

Al examinar por tareas, en relación con las “Elecciones 2021” se observa una mayor concentración del anclaje, superando la media de anclaje de las otras tareas (COVID-19 y Noticias Falsas). Por lo tanto, se evidencia una mayor tendencia al anclaje para todos los estudiantes, independientemente del nivel de EMPC reportado. Según esto, el conocimiento previo de la temática de la tarea podría operar en el sesgo de anclaje en tópicos menos conocidos por los participantes, lo que los llevaría a confiar en la información presentada en la consigna. Por otra parte, en el caso de las tareas “COVID-19” y “Noticias falsas”, la capacidad de autorregulación de los estudiantes con mayor EMPC, podría influir sobre la

orientación al logro, lo que refuerza la idea de que el EMPC opera cuando el nivel de conocimiento de la tarea por parte de los participantes es menor.

Al determinar la influencia de las variables saturación y el EMPC sobre el anclaje, se observó una diferencia significativa respecto de la saturación ($p=1.5 \times 10^{-2}$). De la misma manera, el sesgo de anclaje presenta diferencias significativas respecto del contexto de la tarea. Los estudiantes tienen niveles de anclaje respecto de la consigna de la tarea significativamente diferentes cuando se encuentran saturados de información y en función del tópico de la tarea. Es decir, cuando se enfrentaron la tarea de “Elecciones 2021”, los estudiantes tuvieron una fuerte tendencia a anclarse a los términos propuestos en la consigna de la tarea, tales como “elecciones”, “abril”, “2021”, etc.; mientras que en la tarea “COVID-19” tendieron a utilizar mayor diversidad de palabras tales como “muertes”, “defunciones”, entre otras. Lo que evidencia la influencia del tópico de tarea sobre la selección de términos para las consultas.

Selección de las fuentes de información:

Para estudiar la ocurrencia del sesgo de anclaje en la selección de fuentes de información se calcula el número de palabras contenidas en la relación “consulta - título” y en “consulta - *snippet*” de las fuentes guardadas en la biblioteca del estudiante.

Tabla 8 : Sesgo de anclaje entre la consulta- título de la biblioteca

Dimensión del EMPC		Covid-19		Elecciones 2021		Noticias falsas	
		Experim.	Control	Experim.	Control	Experim.	Control
Expectativa	Alto	0,42	0,37	0,35	0,63	0,44	0,44
	Bajo	0,27	0,29	0,4	0,58	0,55	0,59
Importancia	Alto	0,28	0,29	0,35	0,59	0,54	0,55
	Bajo	0,3	0,33	0,56	0,59	0,5	0,67
Utilidad	Alto	0,28	0,28	0,38	0,54	0,5	0,51
	Bajo	0,32	0,36	0,43	0,7	0,67	0,72
Interés	Alto	0,33	0,26	0,39	0,61	0,5	0,63
	Bajo	0,23	0,35	0,39	0,56	0,57	0,49
Costo	Alto	0,42	0,22	0,33	0,56	0,55	0,77
	Bajo	0,24	0,32	0,41	0,6	0,53	0,5

Fuente: Elaboración propia

En la selección de las fuentes de información respecto del título (Tabla 8) se observan diferencias significativas en la operación del sesgo de anclaje en las distintas tareas. Se observa que el anclaje con el título es menor para la tarea “COVID-19” con un máximo de 0,42. Sin embargo, en las siguientes tareas el anclaje aumenta, con un máximo de 0,7 en la tarea “Elecciones 2021” y un 0,77 en la tarea “Noticias Falsas”. En la tarea “COVID-19” el anclaje con el título es menor en el grupo de control para los estudiantes con alto EMPC. Por el contrario, en las tareas “Elecciones 2021” y “Noticias Falsas” aumenta el anclaje con el título alcanzando la mediana más alta en la tarea “Elecciones 2021” en los estudiantes con

alto EMPC. No se han observado diferencias significativas en el sesgo de anclaje durante la selección de información respecto de la saturación y el EMPC. Sin embargo, se observaron diferencias significativas respecto del contexto de la tarea ($p=2.93 \times 10^{-4}$). Por lo que, los estudiantes evidencian diferencias significativas en el sesgo de anclaje con los títulos al seleccionar las fuentes de información respecto de cada contexto de búsqueda.

En la relación “consulta - *snippet*” de los resultados guardados en la biblioteca es posible establecer si en el estudiante opera en mayor medida el sesgo de anclaje con el título o con el contenido del *snippet*.

Tabla 9: Sesgo de anclaje entre la consulta - snippet de la biblioteca

Dimensión del EMPC		Covid-19		Elecciones 2021		Noticias falsas	
		Experim.	Control	Experim.	Control	Experim.	Control
Expectativa	Alto	0,45	0,46	0,55	0,78	0,2	0
	Bajo	0,4	0,48	0,54	0,58	0,35	0,36
Importancia	Alto	0,39	0,46	0,5	0,66	0,32	0,31
	Bajo	0,4	0,51	0,73	0,49	0,29	0,42
Utilidad	Alto	0,43	0,48	0,56	0,55	0,31	0,15
	Bajo	0,28	0,46	0,43	0,71	0,33	0,75
Interés	Alto	0,48	0,4	0,57	0,63	0,37	0,34
	Bajo	0,32	0,56	0,5	0,57	0,24	0,28
Costo	Alto	0,45	0,27	0,47	0,48	0,6	0,19
	Bajo	0,37	0,52	0,57	0,64	0,2	0,36

Fuente: Elaboración propia

Entre el grupo experimental y el grupo de control (Tabla 9) se observan diferencias de distribución en todas las tareas y entre los estudiantes con alto y bajo EMPC. Según esto, los estudiantes se ven mayormente influenciados en la selección de las fuentes de información por el título del resultado que por sobre el contenido del *snippet*. Al determinar la influencia de las variables saturación y el EMPC, no se observan diferencias significativas respecto de estas variables. Sin embargo, sí podemos evidenciar diferencias en el anclaje respecto del contexto de la tarea ($p=6.1 \times 10^{-5}$). Por lo tanto, los estudiantes presentan diferencias en sus anclajes solo en función del contexto de la búsqueda. Estos hallazgos validan que la selección de las fuentes de información de la lista de resultados tiene como principal criterio el título de la página web.

Construcción de una respuesta a la tarea:

Para determinar la influencia de los resultados de búsqueda sobre la elaboración de respuestas por parte de los participantes analizamos las relaciones “resumen - *snippets* en biblioteca”. En esta relación (Tabla 10) se observan diferencias destacables en la distribución del sesgo de anclaje entre ambos grupos. El grupo experimental presenta menor nivel de anclaje que el grupo de control, sobre todo entre los estudiantes con bajos resultados en la escala EMPC. Según esto, los estudiantes expuestos a una lista de resultados saturada tienden a desvincularse de la información contenida en las páginas webs guardadas en la biblioteca para la construcción de sus respuestas, lo que se acentúa entre los estudiantes con bajo EMPC. Al determinar la influencia de las variables saturación y el EMPC, se observaron diferencias significativas en el sesgo de anclaje sobre los resultados seleccionados respecto del nivel de EMPC ($p=1.87 \times 10^{-17}$). Además, podemos evidenciar diferencias significativas en este anclaje respecto del contexto de la tarea ($p=1 \times 10^{-17}$). Por lo que, tanto el nivel de EMPC como el contexto de la tarea resultan influyentes en la operación del sesgo de anclaje en la elaboración del resumen respecto de los resultados.

Tabla 10: Sesgo de anclaje entre el resumen - snippets de la biblioteca

Dimensión del EMPC		Covid-19		Elecciones 2021		Noticias falsas	
		Experim.	Control	Experim.	Control	Experim.	Control
Expectativa	Alto	0,36	0,22	0,17	0,33	0,2	0,11
	Bajo	0,23	0,31	0,2	0,31	0,07	0,12
Importancia	Alto	0,23	0,32	0,2	0,34	0,11	0,11
	Bajo	0,28	0,22	0,19	0,26	0,05	0,18
Utilidad	Alto	0,24	0,33	0,19	0,32	0,1	0,14
	Bajo	0,26	0,2	0,26	0,29	0,06	0,08
Interés	Alto	0,3	0,35	0,24	0,36	0,13	0,14
	Bajo	0,2	0,24	0,13	0,25	0,05	0,09
Costo	Alto	0,23	0,4	0,23	0,34	0,08	0,29
	Bajo	0,25	0,27	0,19	0,31	0,1	0,08

Fuente: Elaboración propia

Sesgo del algoritmo por la consulta del usuario:

En esta investigación observamos la operación del sesgo de anclaje, a partir de las decisiones del algoritmo de búsqueda. Tal como hemos explicado al inicio de este apartado, el algoritmo del motor de búsqueda se nutre de las palabras y/o términos de las consultas de los usuarios para construir su lista de resultados a partir de medidas de similitud entre estos dos elementos. En este sentido, el motor de búsqueda presenta para cada resultado el título del sitio web y su *snippet*, los que ha seleccionado en función de los términos utilizados en

las consultas emitidas. Por lo tanto, ambas piezas de información (título y *snippet*) resultan relevantes en las decisiones de exploración, selección y parada de los estudiantes.

Tabla 11: Sesgo de anclaje entre consulta - título

Dimensión del EMPC		Covid-19		Elecciones 2021		Noticias falsas	
		Experim.	Control	Experim.	Control	Experim.	Control
Expectativa	Alto	0,31	0,35	0,4	0,5	0,42	0,32
	Bajo	0,29	0,31	0,35	0,36	0,45	0,4
Importancia	Alto	0,3	0,32	0,34	0,41	0,38	0,42
	Bajo	0,29	0,33	0,38	0,3	0,54	0,26
Utilidad	Alto	0,29	0,3	0,37	0,38	0,47	0,37
	Bajo	0,31	0,3	0,31	0,38	0,4	0,47
Interés	Alto	0,32	0,3	0,4	0,43	0,43	0,35
	Bajo	0,27	0,34	0,33	0,33	0,47	0,44
Costo	Alto	0,34	0,28	0,44	0,38	0,45	0,31
	Bajo	0,28	0,34	0,35	0,38	0,45	0,42

Fuente: Elaboración propia

Al observar la distribución del sesgo de anclaje entre “consulta - título” en ambos grupos (Tabla 11) observamos una distribución del sesgo similar para las tareas “COVID-19” y “Noticias falsas”. Sin embargo, existen diferencias de distribución significativas en la tarea “Elecciones 2021”, en ambos grupos entre los estudiantes con alto y bajo EMPC. Al determinar la influencia de las variables saturación y el EMPC, no se observan diferencias significativas en esta operación del sesgo de anclaje respecto del nivel de EMPC ni de la

saturación. Sin embargo, se observan diferencias significativas respecto del contexto de la tarea ($p=2.26 \times 10^{-4}$).

Al estudiar el sesgo de anclaje en la relación “consulta - *snippet*”, observamos que el algoritmo entrega *snippets* con una distribución similar en todas las tareas en los grupos de control y experimental, sin diferencias significativas entre los estudiantes con alto y bajo EMPC (Tabla 12). Esto es técnicamente correcto, debido a que la función de un motor de búsqueda es reconocer aquellos documentos relevantes y similares a la consulta.

Tabla 12: Sesgo de anclaje entre la consulta - snippet

Dimensión del EMPC		Covid-19		Elecciones 2021		Noticias falsas	
		Experim.	Control	Experim.	Control	Experim.	Control
Expectativa	Alto	0,47	0,47	0,5	0,69	0,59	0,36
	Bajo	0,38	0,43	0,5	0,55	0,52	0,49
Importancia	Alto	0,42	0,44	0,45	0,6	0,5	0,49
	Bajo	0,37	0,45	0,56	0,5	0,56	0,4
Utilidad	Alto	0,41	0,45	0,51	0,57	0,53	0,43
	Bajo	0,33	0,41	0,42	0,56	0,5	0,61
Interés	Alto	0,45	0,42	0,51	0,56	0,51	0,46
	Bajo	0,37	0,47	0,49	0,58	0,54	0,5
Costo	Alto	0,46	0,39	0,55	0,52	0,52	0,39
	Bajo	0,39	0,47	0,49	0,59	0,53	0,51

Fuente: Elaboración propia

En la Tabla 12 se observa que, en general, se encontraron palabras regularmente utilizadas en las consultas al interior de cada *snippet* entregado por el motor de búsqueda. En consecuencia, observamos que los términos utilizados por los estudiantes se encuentran dentro de los *snippets* presentados por el motor de búsqueda, principalmente en las tareas “Elecciones 2021” y “Noticias falsas”; lo que podría estar asociado con el tipo de tarea y la disponibilidad de información en cada caso. Al determinar la influencia de las variables saturación y el EMPC, no se observaron diferencias significativas en este sesgo de anclaje respecto del nivel de EMPC. Sin embargo, la saturación ($p=8.92 \times 10^{-17}$) y el contexto de la tarea ($p=3 \times 10^{-18}$) producen diferencias significativas sobre este sesgo de anclaje.

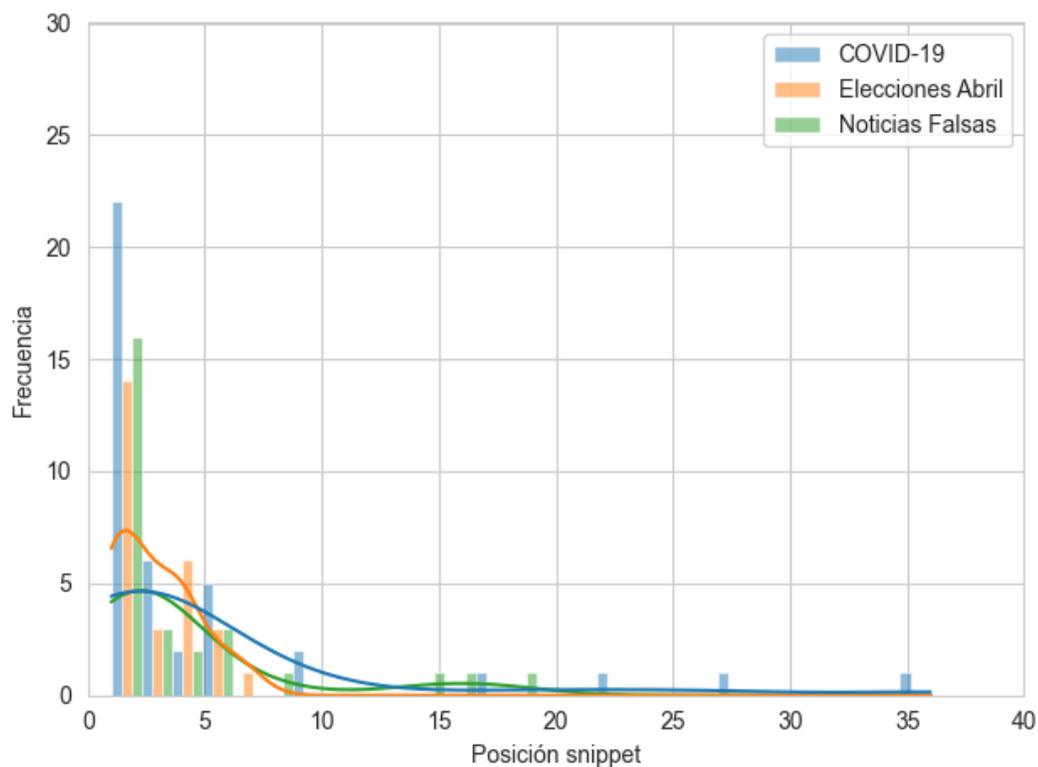
4.2 Sesgo de ranking

El sesgo de ranking corresponde a la creencia extendida entre las personas que los primeros resultados provistos por el buscador tienen mayor relación con la consulta emitida y que se entiende como un comportamiento general de los usuarios (Noble, 2018). En otras investigaciones se ha establecido que las personas seleccionan los resultados que se ubican en las primeras posiciones (Wu y Kelly, 2015; Schultheiß et al., 2018; Su et al., 2018) y son influenciadas por el orden de los resultados propuesto por el motor de búsqueda (Novin y Meyers, 2017). Según esto, el sesgo operaría en los estudiantes al seleccionar los primeros resultados, basados en la confianza en la priorización de resultados propuesta por el algoritmo si el título y/o resumen contienen alguna de las palabras de su consulta.

4.2.1 Caracterización del sesgo de ranking

Dado lo anterior, establecemos que el sesgo de ranking está determinado por la posición del *snippet* de la página web guardada en la biblioteca personal del estudiante en el listado de resultados del buscador. En la ilustración 18, el eje “x” representa la frecuencia de *snippets* y el eje “y” la posición de cada *snippet* guardada en la biblioteca personal del estudiante.

Ilustración 18: Distribución de la posición de la snippet de la biblioteca personal



Fuente: Elaboración propia

Para la tarea “COVID-19” la distribución de la posición de los *snippets*, si bien se concentra en los primeros resultados, se extiende a lo largo del eje “y” (Ilustración 18). Lo anterior, muestra el desacuerdo de algunos estudiantes, respecto de los resultados alcanzados mediante sus consultas, lo que se refleja en un aumento en la profundización de la web. En las entrevistas los estudiantes manifiestan que en las tareas factuales detienen la búsqueda cuando encuentran una respuesta específica, no datos generales. Por lo tanto, la paralización temprana de la búsqueda parece depender directamente del conocimiento previo de fuentes de información válidas. En cambio, para las tareas “Elecciones 2021” y “Noticias falsas” la distribución de la posición de las *snippets* seleccionadas se concentra en los primeros lugares del ranking de resultados (Ilustración 18). En ambas tareas se observa una mayor operación del sesgo de ranking. Donde, la mayoría de las *snippets* guardadas en la biblioteca personal del estudiante se concentran en la primera posición del ranking de resultados. Estos resultados serían congruentes con la literatura. Cabe mencionar que, en las instrucciones

dadas a los estudiantes en la sesión experimental, se les solicita que utilicen esta funcionalidad (guardar en biblioteca) para recopilar aquellas fuentes de información que consideren relevantes para sus resúmenes. Por lo tanto, los documentos guardados en la biblioteca representan fuentes de información que los estudiantes consideraron relevantes de revisar para la construcción de sus respuestas.

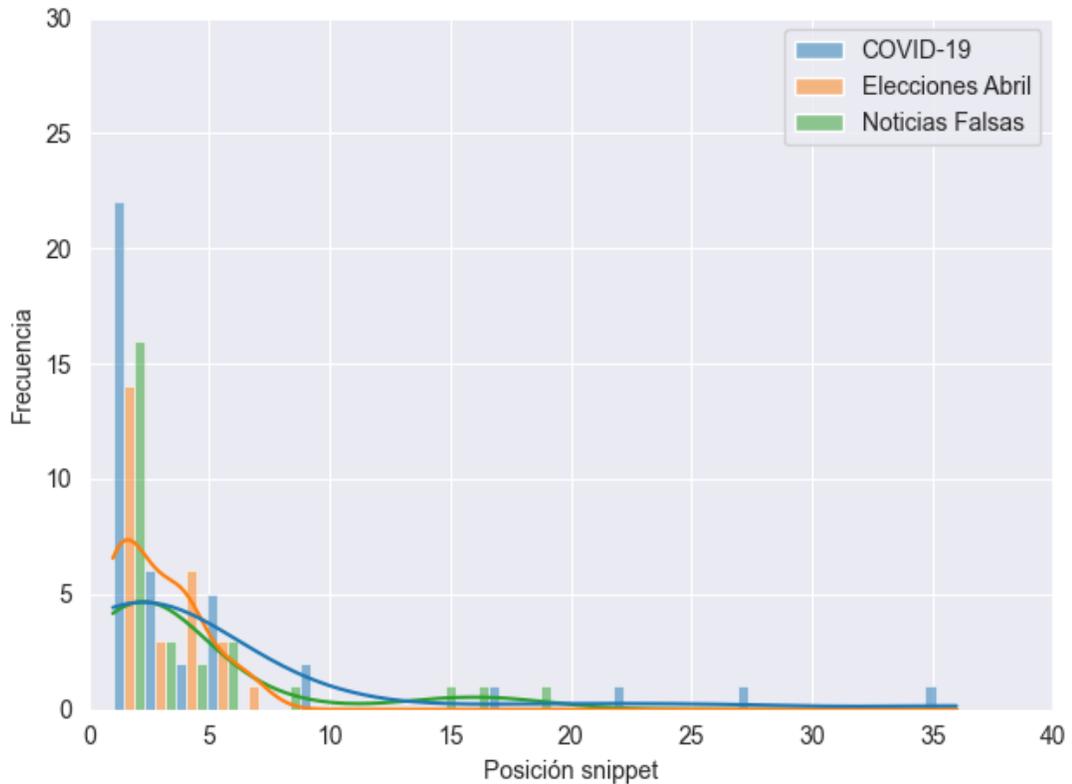
4.3 Sesgo de disponibilidad de resultados

De acuerdo con la literatura, las personas aceptan como respuesta válida la que más se repite entre los resultados de búsqueda (Ghenai et al., 2020). Esto se confirma, ya que los participantes tienden a confiar en los resultados del buscador cuando las fuentes de información son similares entre ellas, tanto en el título como en el *snippet* y con los términos utilizados en las consultas que elaboraron (Tablas 11 y 12). Por lo anterior, en las búsquedas que obtienen fuentes con títulos o *snippets* similares entre ellas se reduce el número de consultas debido a la confianza en la homogeneidad de los resultados de búsqueda. Por el contrario, en fuentes con títulos o *snippets* diversos, los estudiantes tienden a aumentar el número de consultas y la exploración en la web.

4.3.1 Caracterización del sesgo de disponibilidad de resultados

Para representar el sesgo de disponibilidad utilizamos el gráfico de barra que establece la relación entre la cantidad de consultas por usuario para cada tarea, demostrando la decisión de exploración de los resultados por parte de los estudiantes. En la ilustración 19, el eje lateral izquierdo indica el número total de estudiantes y el eje horizontal la cantidad de consultas.

Ilustración 19: Cantidad de consultas por tarea



Fuente: Elaboración propia

En el caso de la tarea “COVID-19” en la que participaron 64 estudiantes se observa un mayor número de consultas por sobre las otras tareas. Esta tarea muestra que ante resultados diversos y alejados de sus consultas los estudiantes emiten más consultas. Como casos atípicos un estudiante elabora 18 consultas y otro emite 34 consultas (Ilustración 19). Para la tarea “Elecciones 2021”, se observan menos consultas, pero se mantiene la relación con la similitud entre los términos de las consultas y los resultados provistos por el algoritmo. En esta tarea participaron 62 estudiantes, de los cuales 34 realizaron una única consulta (Ilustración 19). El sesgo de disponibilidad se refuerza en la tarea “Noticias falsas”, donde de los 62 estudiantes, 39 emiten una única consulta, con apenas 2 estudiantes que realizaron 6 consultas.

5. Conclusiones

En esta investigación se estudia el comportamiento de búsqueda de estudiantes universitarios chilenos para determinar los posibles sesgos y la influencia de la saturación y la motivación de los estudiantes a pensar críticamente.

Para la primera pregunta de investigación ¿Cómo operan los sesgos cognitivos durante la búsqueda de información para resolver una tarea académica? Se logró observar y caracterizar los sesgos cognitivos de: anclaje, ranking y de disponibilidad. El sesgo de anclaje se caracterizó en cuatro instancias: a) La elaboración y/o reformulación de consultas; b) La selección de las fuentes; c) La construcción de una respuesta a la tarea y; d) Sesgo del algoritmo por la consulta del usuario. Se observó la presencia del sesgo de anclaje en cada una de estas instancias del proceso de búsqueda de información, mientras que la intensidad del sesgo cambia principalmente en función del tópico de la tarea. En cuanto al sesgo algorítmico ante la consulta del usuario, se observa que el mayor nivel de anclaje se alcanza en relación con el título para ambos grupos (control y experimental). Las variables saturación y EMPC no muestran una influencia significativa sobre este anclaje, sino que el nivel de anclaje se determina a partir del tópico de la tarea. De acuerdo con lo anterior, a menor conocimiento previo del tópico de la tarea aumentaría el sesgo de anclaje con el título del resultado. Al analizar en detalle las trazas de los estudiantes, se observa que se anclan en los tópicos usados en la consigna de la tarea, aun cuando existan palabras claves que pudiesen conducirlos a una búsqueda más eficiente. Incluso, varios copian y pegan la pregunta en la caja de búsqueda. Dichos términos ancla son utilizados desde la elaboración de la consulta hasta la construcción de su respuesta. Esto explicaría que los participantes revisen solo los resultados que contengan dichos tópicos en el título.

Respecto al sesgo de ranking, se concluye que los estudiantes están altamente influenciados por la decisión del algoritmo de búsqueda. Por lo que se confirma que las personas tienden a seleccionar los primeros resultados del ranking (Wu y Kelly, 2015; Schultheiß et al., 2018; Su et al., 2018). Esto se observa en todas las tareas, especialmente en las tareas “Elecciones 2021” y “Noticias falsas”, donde la mayoría de las fuentes guardadas en la biblioteca se ubican en la primera y segunda posición del ranking. Por

consiguiente, la valoración de una página web como fuente válida de conocimiento está directamente influenciada por la decisión algorítmica. Esto es particularmente riesgoso en el caso de fuentes de información utilizadas para formar opinión y nuevo conocimiento, como en este experimento.

De acuerdo con la literatura, en la caracterización del sesgo de disponibilidad de información definimos que, si varias de las fuentes proporcionadas en la lista de resultados coinciden en la información, ésta se acepta como válida (Ghenai et al., 2020). En los resultados se observa la operación de este sesgo en las tareas “Elecciones 2021” y “Noticias falsas” evidenciado en un menor número de consultas. En contraste, para la tarea “COVID-19” se advierte una reducción de este sesgo, con un mayor número de consultas y/o reformulaciones, con casos atípicos de 12, 18 y 34 consultas. En esta tarea, la decisión de continuar la búsqueda podría estar en las diferencias de los datos de las fuentes. Esta temática en particular fue objeto de debate en varios países debido a las diferencias de cálculo de defunciones por COVID-19 (Baeza-Yates, 2021; The Economist, 2021). Por lo que los estudiantes recibieron distintas cifras, incluidas las oficiales. Lo anterior, marcaría la tendencia de aumentar el número de consultas en esta tarea, debido a que los participantes intentaron encontrar un consenso en la información. Esta estrategia podría considerarse como un criterio fiable, excepto si la temática de búsqueda es objetivo de campañas de desinformación.

En la segunda pregunta de investigación ¿La motivación de los estudiantes a pensar críticamente influye significativamente sobre la operación de sesgos cognitivos? En general, en la elaboración de consultas se observa un mayor sesgo de anclaje en los estudiantes con resultados más altos en la escala EMPC. Sin embargo, se evidencia un aumento en el nivel de anclaje en la tarea “Elecciones 2021”, lo que estaría influenciado por el bajo conocimiento previo de la tarea. Mientras que las respuestas de los estudiantes con bajo EMPC no estarían directamente vinculadas con los resultados alcanzados en la lista de resultados, sino que corresponden a otras fuentes de información tales como el conocimiento previo.

Para la tercera pregunta de investigación ¿Los estudiantes que reciben una lista de resultados saturada manifiestan un comportamiento de búsqueda más sesgado? Durante la construcción de una respuesta a la tarea se logró determinar que la saturación de resultados en interacción con la escala EMPC, influye de manera significativa sobre la operación de los sesgos cognitivos. Mientras que en la selección de las fuentes de información podemos concluir que, de las fuentes guardadas en la biblioteca, se observan diferencias significativas en los niveles de sesgo de anclaje en las diferentes tareas. Sin embargo, los estudiantes seleccionan las fuentes preferentemente basándose en el título del resultado, por lo que el nivel de anclaje es mayor en esta relación (consulta-título). Sin embargo, se observan diferencias en los niveles de anclaje dependiendo del contexto de la tarea en ambos grupos (control y experimental) por sobre el nivel de saturación. En concreto, los estudiantes que recibieron una lista de resultados saturada presentan un menor nivel de anclaje respecto de quienes recibieron una lista no saturada, pero tienden a confiar más en los resultados entregados en las primeras posiciones y a emitir menos consultas. Este comportamiento podría derivar de la confianza que tiene los usuarios en los motores de búsqueda como curadores de información (Noble, 2018). Lo anterior, implica que los participantes parecen no advertir la saturación de información, debido a que solo revisan los primeros resultados y no exploran la lista completa, lo que conlleva a que no se observen relaciones significativas entre los grupos experimentales.

Otro hallazgo inquietante es que los estudiantes que participan de la investigación no muestran habilidades técnicas de acceso y localización efectiva de información, ubicándose por debajo de lo establecido en la literatura, ya que ninguno utiliza las funciones avanzadas de búsqueda (Yamamoto et al., 2018).

Finalmente, estos resultados nos motivan a continuar nuestra línea de investigación y profundizar el estudio para relevar la importancia de las habilidades de búsqueda y selección cuidadosa de información, sobre todo considerando que su utilidad va más allá de la educación formal. Es más, la selección de información que de garantía de calidad y credibilidad es una competencia clave para la participación democrática y el ejercicio activo de la ciudadanía global.

6. Limitaciones y proyecciones

El presente estudio fue realizado con estudiantes universitarios de primer año; por lo que los participantes pueden representar un comportamiento pasivo en comparación a estudiantes de cursos superiores. Sería desafiante proponer en un futuro estudios longitudinales.

Los participantes debían resolver las tres tareas de búsquedas de información en una sesión de trabajo, lo que podría no ser un comportamiento natural para un proceso de búsqueda de información. Por este motivo sería adecuado diseñar el estudio en varias sesiones de trabajo.

En la investigación se utilizó el motor de búsqueda Bing. No obstante, otros motores de búsqueda como Google o You tienen algoritmos de recuperación de información diferentes que pueden influir sobre las decisiones de exploración y selección de información de los participantes. Por consiguiente, en trabajos futuros se propone profundizar el estudio del comportamiento de los estudiantes en tareas de búsqueda de información por medio de la comparación de los sesgos cognitivos respecto de la utilización de otros buscadores.

Capítulo Cuarto

“BÁSICAMENTE, TODO LO PUEDES ENCONTRAR AHÍ: CREENCIAS, SESGOS Y ESTRATEGIAS DE LOS ESTUDIANTES UNIVERSITARIOS EN LOS MOTORES DE BÚSQUEDA”

1. Resumen

Los ciudadanos digitales estamos cada día más expuestos a información falsa o errónea. Sin embargo, no somos conscientes de las barreras que limitan el acceso a información de calidad. Esta investigación indaga las percepciones de los estudiantes universitarios respecto del funcionamiento de los motores de búsqueda y su comportamiento durante una búsqueda en internet. Se realizaron entrevistas semiestructuradas a 13 estudiantes universitarios que previamente participaron de una sesión de búsqueda experimental. Posteriormente, se analizaron los datos de los reportes del comportamiento de búsqueda con las respuestas de las entrevistas para identificar las creencias, criterios de selección de páginas webs y la operación de sesgos cognitivos en la búsqueda de información.

El principal hallazgo es la observación de los sesgos de familiaridad, ranking, confirmación e información. Particularmente, se establece que la preferencia del motor de búsqueda Google está asociada a la operación del sesgo de familiaridad con la plataforma. Esta relación de cercanía con el motor de búsqueda tiene su origen en la integración con Chrome. Asimismo, este sesgo se vincula con el sesgo de ranking, el que se manifiesta en la priorización de la selección de los primeros resultados. De esta manera, la mayoría de los participantes elige los primeros resultados basado en la relación de confianza que han construido en el tiempo con el buscador. Lo anterior, conlleva un bajo número de consultas, independiente del tipo de tarea a resolver.

En consecuencia, los estudiantes no manifiestan cuestionamientos a la selección del algoritmo. En su imaginario, la lista de resultados posiciona los resultados de acuerdo con: la coincidencia con la consulta, calidad de la información e incluso la relacionan con su veracidad. De allí que depositen una confianza casi absoluta en sus resultados.

2. Introducción

La información disponible en la web crece de manera exponencial gracias a las contribuciones de millones de usuarios alrededor del mundo. Esto ha significado una revolución en el acceso rápido a nuevos conocimientos dispuestos en diferentes sitios web. Sin embargo, la calidad de la información ha disminuido con el tiempo debido a diversos factores tales como las campañas de desinformación o por la falta de experiencia de los usuarios que crean contenidos poco fiables. Sin embargo, los usuarios que acceden a los motores de búsqueda de información para encontrar los datos que necesitan no siempre son conscientes de las barreras personales o algorítmicas presentes en este proceso, pues la búsqueda de información en internet es un proceso personal mediado por la tecnología. Es decir, es un proceso complejo porque combina variables humanas (Kuhlthau, 1991) y técnicas que dificultan el acceso a la información.

Las personas accedemos a la información a partir de nuestras referencias culturales, actitudes y percepciones, involucrando también nuestras emociones, lo que condiciona nuestras estrategias de búsqueda. Por lo tanto, a pesar de ser una actividad habitual entre los estudiantes (Brandão et al., 2020), en la literatura se ha descrito que la experiencia de búsqueda puede ser exitosa o frustrante según las habilidades de acceso a la información. Eisenberg y Berkowitz (2000) establecieron que las habilidades del pensamiento crítico son claves para un proceso de búsqueda exitoso porque permiten definir el problema de información y elegir las estrategias pertinentes. Lo que coincide con el estudio de Breakstone y sus colegas (2021), quienes demostraron que prácticas de evaluación crítica, como la lectura lateral, mejoran la capacidad de los estudiantes para seleccionar y explorar información en la web.

Desde la perspectiva del investigador, Vilches (2011) expone que esta diversidad, abundancia y heterogeneidad de la información en internet complejiza también la investigación científica porque dificulta el proceso de sistematización y la fiabilidad de la información. En esta línea advierte que no es suficiente “clicar para saber” porque, aunque la tarea de buscar información en un motor de búsqueda parece simple, “Se necesita un aprendizaje de la interrogación y habilidades específicas para encontrar la información validada”, vale decir, es el investigador quién debe conducir al algoritmo y no al revés.

A pesar de estas dificultades, los motores de búsqueda suman adeptos desde edades tempranas porque su oferta de acceso a información crece en volumen y aumenta en velocidad. Es precisamente este último atributo el que ha contribuido a mejorar la experiencia de búsqueda de los usuarios con la reducción del tiempo de respuesta. Pero, lo que desconoce la mayoría es que la velocidad de respuesta se ha logrado a costo de reducir el espacio de búsqueda (Pariser, 2017). De esta manera los algoritmos de filtro restringen los resultados de búsqueda y los ordenan según el perfil previamente definido (Yoganarasimhan, 2020).

Como corolario, este filtrado algorítmico ha exacerbado los sesgos presentes en la web (Baeza-Yates, 2018). Esto implicaría que las personas estamos expuestas a una sobreabundancia de información parcial, basada en nuestro historial, la geolocalización, idioma, entre otros (Jiang, 2014a; 2014b; Lewandowski, 2017; Cano-Oron, 2019). De esta manera, el buscador nos oferta información de los sitios webs que consultamos frecuentemente o similares, por lo que difícilmente encontraremos información que nos incomode o desafíe nuestras creencias, ya que la lista de resultados se organiza según los datos que el algoritmo ha registrado de nuestras acciones. En este sentido, estos nuevos modelos de usuario, en rigor no han contribuido a mejorar las búsquedas de información porque no impactan en el desarrollo de competencias de alfabetización mediática e informacional. Sino que han permitido facilitar el funcionamiento técnico de los motores de búsqueda y, por consiguiente, aumentar su popularidad entre los usuarios.

No obstante, no podemos traspasar toda la responsabilidad a los motores de búsqueda porque, finalmente, la web se compone de la interacción de personas con objetos culturales, de ahí la concurrencia de sesgos cognitivos. En este sentido, coincidimos con el planteamiento de Baeza-Yates (2018) respecto de la necesidad de relevar la existencia de sesgos en la web para aportar en la toma de conciencia de la existencia inherente de los sesgos. De esta manera, es posible fomentar la reflexión de las personas respecto del proceso de búsqueda de información en internet.

En anteriores investigaciones se han estudiado las principales estrategias de los estudiantes para realizar búsquedas de información. Nagel y sus colegas (2020) analizaron las estrategias de búsqueda en estudiantes universitarios y su relación con el análisis crítico de la información por medio del marco de trabajo de razonamiento crítico en línea, determinado por la existencia de una relación entre las características de los sitios web

visitados y su desempeño crítico. No obstante, los estudiantes tienden a preferir fuentes como Wikipedia por sobre sitios científicos. De la misma forma, Banerjee y sus colegas (2020) estudiaron la relación entre las habilidades de navegación en la web, el comportamiento de búsqueda y las habilidades de razonamiento crítico en estudiantes universitarios en contexto de economía y medicina. Los participantes debían responder una evaluación acerca de sus capacidades de analizar críticamente la información en la web y buscar información para elaborar un ensayo respecto de la tarea entregada. En el estudio se observaron tres patrones de búsqueda asociados a la capacidad de análisis de la información revisada: “Fact-checking” (3%), “Análisis de perspectiva sin Fact-Checking” (41%) y “Evaluación de la credibilidad de la web” (56%). La mayoría de los participantes tuvieron un comportamiento asociado con la evaluación de la credibilidad de la fuente de información, incluso antes de responder o analizar la pregunta entregada. Mientras que solo un grupo minoritario (3%) utilizó mecanismos directos de Fact-Checking para evaluar la calidad de la información en la tarea. Estos análisis han permitido caracterizar el esfuerzo invertido por las personas y cuantificar el impacto de características cognitivas, procedimentales y afectivas sobre el proceso de búsqueda.

En la sesión de búsqueda experimental realizada previamente encontramos evidencias de la operación de los sesgos cognitivos de anclaje, ranking y disponibilidad. En esta investigación nos interesa profundizar los hallazgos de la caracterización automática del comportamiento de búsqueda con las entrevistas a los participantes respecto de sus percepciones y estrategias de búsqueda como medida de contrapeso a las evidencias del proceso experimental. En este sentido, la plataforma de búsqueda de información GoNSA2 (Olivares-Rodríguez et al., 2017), como dispositivo pedagógico, permite visibilizar el comportamiento de búsqueda de los participantes y sus competencias críticas de selección y uso de la información para facilitar la reflexión de los estudiantes respecto de las implicancias de los modelos algorítmicos. El análisis de contraste de los datos empíricos con los testimonios de los participantes busca identificar las percepciones, los criterios de selección de páginas webs y, a la vez, reconocer la operación de sesgos cognitivos en el proceso de búsqueda, tanto los observados en el proceso experimental como los que emerjan durante las entrevistas.

Concretamente, nuestro estudio busca responder:

RQ1: ¿Cuáles son las rutinas de búsqueda en internet que reconocen los estudiantes universitarios?

RQ2: ¿Cuáles son las percepciones de los estudiantes universitarios respecto de los motores de búsqueda en internet?

RQ3: ¿Qué sesgos cognitivos se pueden identificar en la búsqueda de información en internet por parte de estudiantes universitarios?

RQ4: ¿Cuál es la relación de las percepciones de las rutinas de búsqueda con su motivación a pensar críticamente y su comportamiento observado entre los estudiantes universitarios?

4. Metodología

Esta investigación es parte de una investigación de diseño mixto secuencial (Creswell, 2013). El propósito es establecer la coherencia entre la operación de sesgos cognitivos identificados en el comportamiento de búsqueda de información en la web y la reflexión reportada por los participantes acerca de sus decisiones de inclusión y exclusión de documentos. Esta investigación da cuenta de los resultados de las entrevistas y la triangulación de los datos empíricos de las sesiones de búsqueda con los datos cualitativos recogidos en las entrevistas semiestructuradas.

4.1 Participantes

Para la recolección de los datos cualitativos se eligió la entrevista semiestructurada para obtener una reflexión profunda del problema de investigación, considerando la perspectiva y experiencia de los estudiantes (Sautu, et al., 2005; Flick, 2007). Asimismo, este tipo de entrevista aporta la flexibilidad para realizar preguntas abiertas en las dimensiones previamente establecidas e incorporar preguntas que permitan clarificar dimensiones emergentes.

4.2 Muestra

Se utilizó una muestra no probabilística por conveniencia (Flick, 2007). Participaron 13 estudiantes de primer año de Periodismo e Ingeniería Civil Informática. La ponderación de la muestra se realizó post-estratificación de la muestra en base al número de matriculados de la generación de estudio (Piergiorgio, 2007). En cuanto al género la muestra corresponde a 5 mujeres y 8 hombres, debido a que la matrícula de ingeniería corresponde mayoritariamente al género masculino. Los entrevistados autorizaron el uso de sus datos de manera voluntaria, confidencial, anónima y agregada para los fines de la investigación.

4.3 Procedimiento

Esta investigación se compone de 3 etapas. En la primera etapa los participantes respondieron la Escala de motivación al pensamiento crítico, EMPC (Valenzuela y Nieto, 2008). A continuación, resolvieron tres tareas de búsqueda diseñadas en base a los lineamientos de la taxonomía de Anderson y Krathwohl (Urgo et al., 2019) y sometidas al juicio de tres expertos. Los datos de la sesión de búsqueda se registraron en la plataforma GoNSA2 que interactúa con el motor de búsqueda Bing de Microsoft.

Tabla 13: Tareas de búsqueda

Contexto	Consigna de la tarea	Proceso cognitivo	Tipo de conocimiento
Fallecidos COVID 19	Mencione la cantidad de personas fallecidas por el COVID 19 en Chile durante el año 2020.	Recordar	Factual
Elecciones 2021	Resuma las razones del cambio de fecha de las elecciones de abril del 2021.	Comprender	Exploratorio
Noticias Falsas	Resuma las características de las noticias falsas.	Comprender	Exploratorio

Fuente: Elaboración propia

En la segunda etapa se realizaron entrevistas semiestructuradas, considerando las creencias y experiencia previa de los estudiantes y la posibilidad de levantar categorías emergentes (Vilches, 2011; Flick, 2007; Gaitán y Piñuel, 1997). Las entrevistas se realizaron entre el 22 y el 26 de abril de 2021. Dichas entrevistas fueron grabadas y posteriormente

transcritas para analizar su contenido. Se utilizó la transcripción literal, incluyendo los posibles errores gramaticales y/o reiteraciones para evitar influir en la interpretación de las respuestas. Cada tema fue indagado con una pregunta principal y fue complementado por preguntas secundarias (Tabla 13). De las entrevistas surgieron tres nuevos temas que fueron indagados a partir de nuevas preguntas, tal como se detalla en la tabla 2. Las preguntas se realizaron en un lenguaje directo y llano, realizando las aclaraciones que fueren necesarias para la comprensión de estas. Asimismo, se dio el tiempo necesario para que el estudiante explique con sus propias palabras sus creencias respecto del uso de los motores de búsqueda. En los casos que fue necesario, se les pidió aclarar sus respuestas para evitar errores de interpretación.

En la tercera etapa se desarrolló la triangulación de los datos para establecer las posibles relaciones entre las trazas del proceso de búsqueda, la escala EMPC y las respuestas de los participantes en las entrevistas.

Tabla 14: Categorías preestablecidas

Tema	Preguntas
Rutinas de uso	¿Cuántas horas crees que lo usas durante en el día? ¿Qué otros usos le das al buscador?
Preferencia de motor de búsqueda	¿Cuál es tu favorito? ¿Cuáles otros has utilizado?
Criterios de inclusión/exclusión de páginas webs	¿Cómo decides qué documento abrir? ¿Cómo seleccionas los documentos? ¿Los criterios de selección de las páginas cambian si la búsqueda es para una tarea o es un tema que te interesa?
Elaboración de consultas	¿Cómo decides los términos o palabras que vas a usar para realizar la búsqueda? ¿Te ha ocurrido que inicias una búsqueda y no encuentras ningún resultado satisfactorio?
Habilidades de búsqueda de información	¿Cuál es tu percepción respecto de tus habilidades de búsqueda de información en internet? ¿Qué habilidades crees que te faltan?

Criterios de parada	¿Cuál es el principal criterio que utilizas para detener la búsqueda de información? ¿Cómo estás seguro de que la información que encontraste es útil?
---------------------	--

Fuente: Elaboración propia

Tabla 15: Categorías emergentes

Tema	Descripción	Pregunta
Idioma	Uso un idioma distinto al nativo	¿Por qué buscas en inglés?
Valoración	Percepción de utilidad en su vida cotidiana	¿Cómo valoras la utilidad del motor de búsqueda? ¿Qué crees que pasaría si no existieran?
Ocio y/o entretenimiento	Uso en actividades de ocio y/o entretenimiento	¿Qué otros usos le das al buscador?
Profundización de navegación	Revisión de las listas de resultados	¿Alguna vez has pasado a la página 2 o 3 del listado de resultados? Cuando haces una búsqueda de tu interés ¿avanzas más en las páginas o avanzas más si es con nota?

Fuente: Elaboración propia

5. Resultados

¿Cuáles son las rutinas de búsqueda en internet que reconocen los estudiantes universitarios?

En el primer acercamiento con los entrevistados le preguntamos si utilizaban un motor de búsqueda y en caso de hacerlo cuánto tiempo creían que lo usaban durante el día. Sin embargo, en varios casos los estudiantes no conocían el término y fue necesario aclarar que es un motor de búsqueda. De sus respuestas se infiere que muchos confunden el motor de búsqueda con el navegador.

Estudiante 9: “Porque yo uso uno en específico que es Opera GFX creo, no GX y ese me deja como limitar los procesos, limitar la RAM que usa el buscador o el CPU que usa el buscador.”

No obstante, aclarada la pregunta, una parte minoritaria de los entrevistados insiste en la idea de que no utilizan motores de búsqueda. Por ejemplo, un estudiante confunde el navegador con el buscador y asume que no necesita la mediación de un motor de búsqueda. La percepción del tiempo de uso diario del buscador varía entre 1 hora a 10 horas como máximo. En general, los entrevistados creen que su uso aumenta con la carga académica. En las respuestas se observa que el motor de búsqueda es una herramienta útil en el trabajo académico.

Estudiante 13: “Pues casi todo el día porque tengo muchas tareas y porque estoy estudiando inglés en una página.”

Al preguntar respecto de los usos del buscador, se observa que los estudiantes lo asocian preferentemente al uso académico. Sin embargo, también reportan su uso en consultas de ocio y/entretenimiento.

Estudiante 8: “Eh, pa' hacer tareas mayormente. Para buscar, nos piden así, ya busquen tal documento y tienen que leérselo y ahí ya lo busco en Google. O sea, pa' buscar algo, me quiero descargar algo lo busco por ahí.”

Estudiante 12: “Siempre como que quiero saber cosas, por ejemplo, no siempre tiene que ser como alguna tarea o algo así, hasta algo como normal. Sí, cómo, no sé ayer, lo de los Oscar, para saber quiénes estaban nominados. Cada vez que tengo una curiosidad me pongo a mirarlo para saber la información completa. En ese sentido soy bien curiosa y me meto a Google.”

Si bien, reportan que lo usan principalmente para buscar información académica, también declaran utilizar su buscador favorito para encontrar respuestas que completen vacíos de información de todo tipo. Incluso, suelen utilizarlo como plataforma de entretenimiento.

En sus respuestas observamos que los estudiantes manifestaban la utilidad del buscador en su trabajo académico y también a todo tipo de necesidades de información. Por lo que indagamos respecto de la valoración de los entrevistados a este tipo de plataformas.

Estudiante 13: “Pues creo que es muy útil porque básicamente todo lo puedes encontrar ahí.”

Estudiante 6: “Es casi indispensable diría yo, o sea ahora al estilo de vida al que estoy acostumbrado es casi indispensable. Porque la información la entrega muy fácil, no sé qué haría, cómo la buscaría y todo lo que demoraría sin el motor de búsqueda.”

Todos los estudiantes declaran que valoran positivamente la utilidad del motor de búsqueda en su vida cotidiana. Principalmente aprecian: la accesibilidad, la facilidad de uso, la velocidad y el volumen de la información indexada.

¿Cuáles son las percepciones de los estudiantes universitarios respecto de los motores de búsqueda en internet?

En este sentido, los situamos en un escenario hipotético de inexistencia de los motores de búsqueda y las respuestas sorprenden por su honestidad. La mayoría considera que no sabría cómo acceder a la información, debido principalmente a su abundancia y a que desconocen cuáles son las fuentes fiables. Esa decisión siempre a la tomado el algoritmo, por lo que la mayoría de los consultados no se imagina cómo comenzar a buscar.

Estudiante 6: “Es casi indispensable diría yo, o sea ahora al estilo de vida al que estoy acostumbrado es casi indispensable. Porque la información la entrega muy fácil. No sé qué haría, cómo la buscaría y todo lo que demoraría sin el motor de búsqueda.”

Estudiante 12: “¡Ah no! yo encuentro que en esta época es muy útil. No sabría cómo manejarme sin tener un buscador, ya es costumbre y te trae la información a la mano.”

En el imaginario estudiantil, la inexistencia de la curatoría del buscador nos conduciría de regreso a tiempos enciclopédicos con búsquedas de información lentas, confusas e incompletas.

Estudiante 8: “O sea, tendríamos así como, habrían diccionarios de enlace. Yo encuentro que habrían diccionarios de enlace, así como las amarillas, pero de puros enlaces. Porque al final igual te puedes conectar a una página sin un motor de búsqueda.”

Como caso atípico el estudiante 8 cree que es posible acceder a la información sin la mediación del algoritmo. Sin embargo, los otros entrevistados no consideraron la posibilidad de acceder de manera directa a las páginas webs, sino que destacan al buscador como la herramienta que les facilita un acceso rápido a la información, por lo que la consideran indispensable para elaborar sus tareas académicas.

En cuanto a las estrategias de elaboración de consultas al motor de búsqueda, los participantes explican que intentan elaborar una consulta acotada. De acuerdo con los entrevistados una búsqueda efectiva requiere elaborar una consulta que facilite acotar los resultados de búsqueda.

Estudiante 8: “Eh, trato de poner lo esencial, muchas veces pongo si es que la palabra. No pongo una frase, una oración. Por ejemplo, en el trabajo que tuvimos que hacer de fallecidos por COVID 19, no pongo cuántos fueron los fallecidos por COVID 19, sino que pongo fallecidos COVID 19 y el año, pongo lo mínimo no más.”

La mayoría de los estudiantes reporta que la estrategia más eficiente es usar el mínimo de palabras o términos. Explican que al ejecutar consultas extensas el buscador tiende a ampliar las posibilidades de fuentes aumentando el tiempo de búsqueda. Incluso, hay respuestas que manifiestan usar pocas palabras como estrategia para clicar más rápido.

Estudiante 9: “Sinceramente, como las menos palabras posibles, lo más corto. Por ejemplo, como un resumen de lo que quiero buscar. Por ejemplo, quiero buscar los

fallecimientos de Chile por el Coronavirus actualmente busco "fallecimientos coronavirus Chile", eso. Es como para menos tiempo, clickear más rápido.”

Sin embargo, solo algunos estudiantes mencionan que identifican las palabras clave del tema de búsqueda como estrategia para alcanzar resultados relacionados con su consulta.

Estudiante 4: “Intento buscar por palabras claves más que nada. Por ejemplo, si me preguntan tiene algo que ver con tres cosas fundamentales como por ejemplo un país, una situación y una fecha, entonces busco de esas tres y ya.”

No obstante, existen estudiantes que reportan el uso de la pregunta como estrategia válida de búsqueda

Estudiante 10: “Usualmente, es como yo haría la pregunta a una persona, como me dirijo a una persona. Ponte tú, si quiero saber, ocupando el ejemplo del calendario de vacunación, pregunto cuándo le toca a mi grupo etario vacunación en Chile y así. Mis amigas me dicen que busco como señora porque ellas ponen la pura información así: vacunación Chile, y yo pongo la pregunta entera (ríe).”

Este participante indica que formula preguntas al buscador como lo haría con un interlocutor, aunque reconoce que es una estrategia arcaica suele utilizarla.

Al preguntarles respecto de sus habilidades de búsqueda de información algunos estudiantes consideran que son efectivos localizando información.

Estudiante 12: “Ah yo encuentro que tengo buenas habilidades. Es que claro, nosotros, por lo menos yo, uno crece usando el buscador, desde chico (dice convencido). Entonces, ya está familiarizado y es fácil de usar.”

En sus respuestas también se observa que, tal como en el caso estudiado por Pereira y sus colegas (2019) los entrevistados aprenden a utilizar el buscador por sus propios medios y sienten que sus habilidades mejoran mediante la práctica constante.

Estudiante 1: “Mmm no sé es que no encuentro que haya que ser muy habilidoso para buscar información en internet, o sea es cosa de escribir lo que uno quiere saber no más. Aunque igual es distinto eso a elegir información (se muestra dudativo).”

Como se observa en la respuesta anterior, varios entrevistados se consideran competentes elaborando una consulta. Pero, esa seguridad para localizar información se reduce al indagar respecto de las competencias para seleccionar y usar la información disponible en la web.

Estudiante 3: “Yo diría que actualmente mis habilidades de búsqueda son bastante pobres (ríe) dado que generalmente hago búsquedas muy relacionadas al día a día y no hago muchas búsquedas específicas o centradas, no muy académicas. Entonces si me piden buscar un dato muy específico de algún tipo de estudio no tengo conocimientos sobre ni buscadores ni cosas así que me puedan ayudar a facilitar esta tarea.”

En general, la mayoría de los entrevistados considera que sus habilidades de búsqueda son bajas o promedio. Varios de ellos se identifican a sí mismos como usuarios pasivos de los buscadores.

Estudiante 9: “(repite en voz alta la pregunta) Mmm (piensa) son simples y soy como, usualmente, supongo que todos, como que se quedan con el primer resultado que encuentran.”

Reconocen, por ejemplo, dificultades para elaborar las consultas y para construir una respuesta propia con la información disponible.

Estudiante 7: “Uy, ahí no más (ríe). porque si es sobre un tema donde, por ejemplo, tengo que hacer yo mi propia respuesta, me cuesta idear una pregunta para buscar la información, para saber de un tema en específico. Porque igual no puedo sacar

información de una página tan poco confiable, tengo que buscar algo con más datos que corroboren que es verdad.”

Entre las carencias también admiten que cuentan con un vocabulario reducido que limita sus posibilidades de hacer un buen uso de los motores de búsqueda.

Estudiante 4: “Conocimiento de más vocabulario creo yo como, por ejemplo, es difícil intentar buscar algo si no tienes la palabra clave necesaria, entonces puedes intentar parafrasear un poco, pero igual te va a tirar muchos resultados que son o pueden ser distintos, entonces creo que un poco más de vocabulario sería importante.”

Sin embargo, su imaginario respecto la operación algorítmica de los motores de búsqueda los lleva a establecer como estrategia la reformulación de la consulta reduciendo el número de palabras, procurando ser más eficientes, cuando intentan encontrar lo que buscan. Esta táctica es eficiente y opera de acuerdo con la lógica del algoritmo, sin embargo, los estudiantes no asientan dicha estrategia basados en el conocimiento del funcionamiento del algoritmo. Sino que su estrategia se origina en el pensamiento intuitivo o heurístico, que en este caso es eficiente y no conduce a un sesgo cognitivo (Matute, 2019).

¿Qué sesgos cognitivos se pueden identificar en la búsqueda de información en internet por parte de estudiantes universitarios?

En una de las primeras preguntas les pedíamos que nos dijeran cuál era su buscador favorito. Todos los estudiantes manifestaron utilizar preferentemente Google.

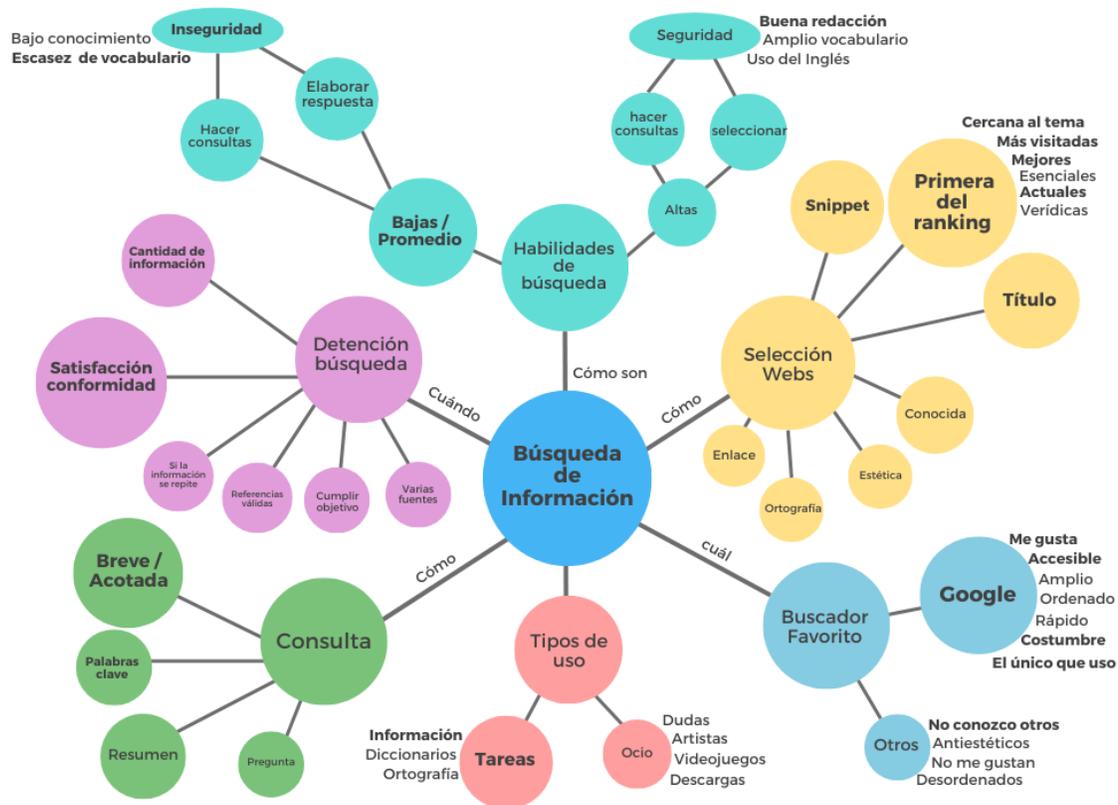
Estudiante 11: “Google, sí. Porque según yo es el que, como que he probado otros, pero Google es como el más rápido, como que muestra la mejor información según yo, al menos esa sensación me da”.

Al profundizar respecto de las razones de la preferencia de Google por sobre otros motores de búsqueda, en general, las respuestas emplean argumentos emocionales, tales como:

Estudiante 2: “Lo siento más amplio, como que puede que no sea muy específico, pero da varias visiones de varias cosas y ahí uno obviamente tiene que aplicar el filtro de las fuentes y que busca principalmente, sobre todo en ese aspecto”.

No existen respuestas enfáticas respecto de la superioridad técnica o de la diversidad de fuentes de información. Sino más bien, los estudiantes reconocen un acostumbramiento debido al uso habitual de Google. En las entrevistas admiten que usar Google es un hábito difícil de cambiar, a pesar de que varios indican conocer otros buscadores, prefieren continuar usando Google. De esta forma, los estudiantes sin saberlo están utilizando el sistema 1 de

Ilustración 20: Mapa de entrevistas semiestructuradas



Fuente: Elaboración propia

razonamiento intuitivo. Ahora bien, las personas constantemente recurrimos a heurísticas para tomar decisiones en ambientes de incertidumbre, es inherente de nuestra condición humana. La complejidad surge cuando usamos la decisión intuitiva en decisiones que requieren un mayor esfuerzo cognitivo.

Estudiante 10: “Es que me gusta Google. A veces he probado otros, pero como que me acostumbre a ese (dice pensativo).”

Una de razones que facilitarían esta tendencia al uso preferente de Google es la configuración estándar de Chrome, la que establece a Google como motor de búsqueda por defecto.

Estudiante 6: “Es el único que uso en realidad porque es accesible porque uso Chrome, porque uso Chrome y Chrome lo trae integrado y, por accesibilidad (se muestra seguro).”

Un grupo importante indica que es el único que conoce. De esta manera, la relación de familiaridad con la plataforma se inicia desde que el estudiante comienza a utilizar su computadora. La familiaridad es una decisión heurística automática que tiene la particularidad de actuar como un estímulo afectivo, lo que facilita sesgar la decisión (Cortada, 2008).

Estudiante 13: “No sé, lo he usado desde siempre y no lo he cambiado porque para mí funciona.”

De acuerdo con la literatura, el sesgo de familiaridad proporciona una sensación de seguridad que predispone la baja atención o de alarma (Matute, 2019). Lo anterior, contribuiría también a sesgar la opinión de los estudiantes respecto de los otros motores de búsqueda que participan del mercado.

Estudiante 8: “Los otros no me gustan, Yahoo o DuckDuckGo, Bing, los encuentro desordenados. Estéticamente no me gustan, es por eso (expresa con desdén). Sí, es que la costumbre de ocupar Google.”

En definitiva, la facilidad técnica de la integración de Google en el sistema operativo y los navegadores web facilitaría la operación del sesgo de familiaridad con este motor de búsqueda, alzándolo como el favorito de los estudiantes y, por ende, con la mayor cuota de mercado.

A continuación, les preguntamos a los participantes cómo seleccionaban los documentos a revisar de la lista de resultados. En sus respuestas nos revelan que solo revisan la primera página de resultados del motor de búsqueda porque confían en las decisiones del algoritmo de Google. Ellos suponen que el listado de resultados está ordenado en base a las páginas webs más cercanas a su consulta, incluso lo relacionan con información verificada.

Estudiante 8: “Porque encuentro que, en la primera página, yo lo digo por experiencia, por lo menos sale lo que o es verídico, o es esencial, o está confirmado. Porque después lo otro son blogs de gente, páginas de Turquía no sé, que alguien creo en el sótano, cosas así.”

A pesar de la confianza en la recomendación del buscador, los entrevistados declaran desconocer la estrategia del algoritmo de Google para elaborar el ranking de resultados, indicando que su confianza se basa en su experiencia de búsqueda. No obstante, muchos reconocen que podrían estar equivocados en su razonamiento.

Estudiante 1: “Porque las primeras páginas son las mejores según yo. No sé, supongo que son las que la gente más visita, aunque en verdad no sé porque te muestra unas antes que otras, o sea yo supongo eso.”

Al intentar profundizar respecto de sus razones para no revisar más allá de la primera página, los estudiantes explican que no suelen hacerlo porque lo consideran infructuoso y reafirman convencidos que los primeros resultados son los mejores.

Estudiante 2: “Si no encuentro en la primera, normalmente no llegaría a la segunda porque en general el algoritmo de Google tiende a dejar en primera instancia lo más

actual, lo más reciente, donde las personas más acceden, tiende Google a subirlas en sus categorías, tienden a ser las primeras en las páginas.”

En definitiva, esta simplificación de las decisiones de búsqueda que ofrece el motor de búsqueda ha propiciado un comportamiento pasivo en los estudiantes (Fernández-Vicente, 2020). En efecto, no se observan prácticas de resistencia durante la búsqueda de información en la plataforma GoNSA2 ni durante las entrevistas. Es más, los participantes afirman delegar la selección de la información al algoritmo porque consideran que reciben la información que necesitan. En sus respuestas los estudiantes manifiestan una actitud acrítica que no cuestiona la parcialidad del listado de resultados (Vidal-Sepúlveda y Olivares-Rodríguez, 2021). Esta confianza se traduce en un sometimiento a la personalización del algoritmo derivado de la operación del sesgo de familiaridad con su buscador favorito.

En esa línea, indagamos respecto de sus criterios de credibilidad de páginas webs. Principalmente, los estudiantes indican priorizar las webs de las primeras posiciones del ranking de resultados.

Estudiante 12: “Generalmente, miro los primeros porque una parte de mi dice si están primeros es porque son mejor, o si no, de acuerdo al título, si se acerca a lo que estoy buscando. A veces, abajito, en el caso de Google dice tiene ciertos términos que buscaste y busco el que tenga todos los términos que yo puse (se observa seguro en su respuesta).”

De acuerdo con lo anterior, los estudiantes tienden a confiar en la fuente de información si se ubica entre los primeros resultados (Wu & Kelly, 2015; Schultheiß et al., 2018; Su et al., 2018; Noble, 2018; Liu & Zhang, 2019). El sesgo de ranking opera basado en la creencia popular de que la selección algorítmica ubica a los mejores resultados en las primeras posiciones.

Estudiante 9: “La primera página, Yo creo que porque la gente va buscando el mejor resultado y ese resultado va escalando hasta llegar al principal, que es al que todos entran o se quedan más tiempo en la página.”

De acuerdo con este estudiante la lista de resultados se elabora basado en la popularidad de las páginas webs, lo que no podría entenderse como un indicador de calidad ni veracidad de la información. Además, en su respuesta se manifiesta el desconocimiento respecto de las estrategias de posicionamiento de las páginas web en los motores de búsqueda (Gonzalo, 2015).

Otros criterios mencionados en las entrevistas son: el título del resultado y las palabras claves destacadas en el resumen del *snippet*.

Estudiante 13: “No sé, los elijo por el título normalmente. A veces, por lo poco que aparece de información que tiene adentro, como que aparece el título y algo que dice.”

También declaran usar páginas webs conocidas. En este caso observamos nuevamente la operación del sesgo de familiaridad (Matute, 2019) esta vez con la fuente de información. Este sesgo operaría si el estudiante selecciona fuentes conocidas o que ha usado previamente por sobre otras fuentes válidas.

Estudiante 4 “Si la página es una página que suelo frecuentar, si es un noticiero que suelo frecuentar le doy clic a ese y eso. Por lo general igual siempre que busco algo la respuesta no pasa de la primera página del buscador, intento nunca sobrepasar de esa página. Porque yo creo, al menos desde mi percepción, creo que el resto de páginas ya empieza a nada. No sabría decir si son muy alejadas del tema, pero páginas que son poco visitadas y creo que eso le resta valor al momento de decidir, al menos para mí. Igual puede que sea todo lo contrario y que de verdad las páginas sean muy buenas, pero yo me baso así más que nada (ríe).”

Otros estudiantes mencionan el enlace y el dominio de la página web como alertas de precaución sobre fuentes con dominios poco conocidos.

Estudiante 5: “(piensa) Por lo general en el enlace el nombre, el nombre del enlace ya dice bastante y el dominio también, o dentro de la página la organización se nota cuando hay dedicación y cuando no. Si, por ejemplo, usted ve a un lado lleno de publicidad,

publicidad tras publicidad, hay algo raro ahí. En cambio, cuando aparecen enlaces directo a su propia página de otra noticia y se ve un formato más profesional se ve que la página es un poco más confiable.”

Asimismo, aluden a la apariencia de la página web, aunque este criterio es posterior, ya que se aplica una vez que se ha seleccionado el resultado.

Estudiante 1 “Dependiendo de la página que los suba, por ejemplo, me sale información de una página que la abro y se ve terrible fea, y se ven poco profesionales, por ejemplo, entonces, ahí esa no la uso (dice convencido).”

En general, los estudiantes indican que la mayoría de las veces quedan conformes con la priorización de la información de Google. Estos resultados serían congruentes con la literatura y la confianza expresada por los estudiantes en las entrevistas respecto de las decisiones algorítmicas. Todos manifiestan que las mejores páginas webs corresponden a la primera página de resultados. Por lo tanto, no buscan en las siguientes páginas, dado que estiman que dichos resultados se alejan de su consulta. Solo dos de los entrevistados indican que las posibilidades de encontrar información aumentan cambiando el idioma al inglés y lo destacan como una estrategia que suelen utilizar.

Estudiante 6: “Eh yo creo que soy bueno buscando cosas en internet, es que como busco todo me acostumbre más o menos con que palabras tengo que buscar. Y, sobre todo, busco en inglés, cuándo no encuentro casi siempre está.”

También estábamos interesados en conocer cómo deciden detener una búsqueda de información, por lo que les preguntamos directamente si tenían algún criterio definido. La mayoría de los estudiantes nos indicó que detienen la búsqueda de información cuando consideran que tienen una cantidad suficiente de información que les permita responder a su investigación inicial. Sin embargo, no tienen una referencia estándar para determinar la cantidad de información necesaria, sino que la definen como una sensación de satisfacción personal con la cantidad y el tipo de datos recogidos.

Estudiante 7: “(piensa) No sé cómo explicarlo, pero yo siento que ya está bien respondido, entonces... Eh, igual la cantidad de información igual es importante, pero creo que es más como un sentimiento dónde digo ah acá nombré una parte importante y está bien.”

Entre los estudiantes que declaran usar como criterio la cantidad de información se observa la operación del sesgo de información, entendido como la tendencia a creer que más información implica más veracidad. Esto porque podría crear la falsa impresión de que, por tener más información, el razonamiento o la conclusión será más veraz (Baron, 2000).

Un pequeño grupo indica que su criterio de parada está vinculado con su objetivo o pregunta inicial, por lo que la cantidad de páginas que revisen es irrelevante. La búsqueda se detiene si consideran que la información se ajusta o no a lo que están buscando. En este caso podríamos considerar que operan mecanismos de autorregulación. Sin embargo, en sus respuestas son enfáticos en declarar que buscan responder a lo que buscan inicialmente.

Estudiante 9: “Cuando me satisface lo que la página me da y cuando tiene todos los datos que yo busco.”

Lo anterior, podría también sugerir una tendencia al sesgo de confirmación, debido a que no manifiestan preocupación por la veracidad sino a la tendencia de favorecer aquella información que confirme sus presupuestos (Matute, 2019).

Otro grupo minoritario declara, como criterio de parada, la revisión de las referencias de la fuente. De este modo si la información está respaldada por referencias que estimen como válidas, aceptan la información y cierran su búsqueda. Sin embargo, varios de ellos declaran que si la búsqueda es de interés personal los criterios de búsqueda se flexibilizan.

Estudiante 6: “Cuando busco información académica sí, eh intento fijarme en las referencias más que nada, y ver por ejemplo el autor, buscar el autor, ver en qué cosas ha trabajado. Porque si es un autor, que busco el autor y me sale, por ejemplo, por decir algo que es antivacunas o algo así, voy a empezar a dudar de su credibilidad en temas

de investigación y cosas así. Pero, en algo que me gusta no siempre busco, no siempre busco la fiabilidad por así decir.”

Como caso atípico un estudiante indica que busca comparar fuentes distintas. Si hay consenso acepta como válida la información y detiene la búsqueda.

Estudiante 4: “Por lo general intento corroborar que la duda que tenía o la tarea que tenía es respondida, pero de varias fuentes, no solo de una, si fuera de una no me sentiría cómodo terminado la búsqueda, intentaría buscar al menos unas dos páginas, o sea tres páginas que me aseguren que la respuesta que tengo es lo más acertada posible. Pero siempre sean tres documentos, estoy muy seguro de eso. Lo estoy pensando ahora y nunca me había detenido a pensar en eso, pero ahora que veo siempre decido tres documentos como mínimo, para asegurarme de que mi respuesta es correcta y para tener la seguridad de que mi respuesta está bien.”

Los estudiantes explican que esperan que una búsqueda les proporcione de manera rápida y eficiente la información que buscan. Lo que se expresaría en: la cantidad de información, el número de ejemplos, la satisfacción personal con los datos y la redacción de la información.

Del análisis de las entrevistas observamos la operación de cuatro sesgos cognitivos: familiaridad, ranking, información y de confirmación. Es importante destacar que los participantes no mencionan en sus respuestas el término “sesgos” o similares, lo que podría indicar desconocimiento de su ocurrencia en la web, lo que refuerza la idea de la necesidad de propiciar la reflexión de los usuarios de los motores de búsqueda respecto de la existencia inherente de sesgos en la web.

¿Cuál es la relación de las percepciones de las rutinas de búsqueda con su motivación a pensar críticamente y su comportamiento observado entre los estudiantes universitarios?

Para profundizar el análisis de las entrevistas contrastamos las percepciones declaradas en sus respuestas con su comportamiento durante el proceso de búsqueda experimental. En esta etapa se consideró a los entrevistados que completaron la totalidad de las tareas a resolver, alcanzando un total de 10 participantes.

a) Tipos de consultas:

Entre los participantes se identifican distintas estrategias para elaborar las consultas (Tablas 16, 17 y 18). La estrategia de elaboración de la consulta no es dependiente de la persona (usuario) sino que muestra variaciones en relación con el conocimiento de dominio de la tarea. De esta manera, a menor conocimiento del dominio los participantes tienden a anclarse en las palabras utilizadas en la consigna de la tarea y, en sus posteriores reformulaciones, la tendencia es alargar la consulta con la incorporación de nuevos datos para especificar la búsqueda. Sin embargo, este resultado es contrario a lo manifestado en las entrevistas, dónde indicaban que siempre buscaban realizar consultas breves y acotadas. Asimismo, se observa que el uso de palabras clave de elaboración propia se relaciona con un mayor conocimiento del dominio de la búsqueda, evidenciado en un menor número de consultas. Estos resultados reafirman las dificultades manifestadas por los estudiantes para buscar información, dado que la mayoría reconoció tener dificultades para elaborar consultas principalmente por escasez de vocabulario.

b) Fuentes seleccionadas:

Los participantes tenían la opción de guardar en un espacio de biblioteca personal aquellas fuentes que consideraran relevantes de revisar. En todas las tareas los estudiantes guardan mayoritariamente páginas webs de medios de comunicación. Asimismo, al analizar las páginas webs utilizadas como información válida para resolver las tareas, los estudiantes usaron preferentemente páginas webs de medios de comunicación, independiente del tipo de tarea a resolver. Además, se observa que principal criterio para seleccionar páginas webs es el ranking del listado de resultados. Así, incluso aquellos participantes con altos resultados en la escala EMPC, son influenciados por la selección algorítmica, tal como lo manifestaron en las entrevistas.

c) Resumen:

De acuerdo con la literatura, la complejidad del producto solicitado aumenta el esfuerzo y las estrategias de exploración de los participantes, produciendo mayores evidencias del desempeño (Olivares-Rodríguez y Guenaga, 2015; Novin y Meyers, 2017). En este caso los estudiantes resolvieron una tarea factual “COVID 19” y dos tareas exploratorias “Elecciones 2021” y “Noticias Falsas”. Por lo que solo estas últimas requerían un esfuerzo mayor de exploración de la web y la creación de un resumen de la información recopilada. Sin embargo, al analizar las respuestas de la tarea “COVID 19” se observó que solo 6 de los 10 participantes entregaron una cifra cercana a la oficial. En cambio, en las tareas exploratorias no se observan errores en las respuestas sino diferentes niveles de esfuerzo en la redacción del resumen. Para la tarea “Elecciones 2021” la mayoría de los estudiantes realiza un resumen de la información revisada, solo uno copia de manera textual una cita a un discurso de Sebastián Piñera (ex presidente de Chile) consignada en el medio. De igual manera en la tarea “Noticias Falsas” la mayor parte de los estudiantes resume la información recopilada, pero aumenta el número de estudiantes que copian y pegan información. No obstante, aumentan los participantes que elaboran resúmenes basados en la investigación de dos o más fuentes.

Análisis trayectorias de búsqueda

A continuación, analizamos la trayectoria de búsqueda de 4 casos representativos contrastando los datos del comportamiento en el experimento y la entrevista. Cabe mencionar que en esta etapa de análisis se consideró a estudiantes de ambos géneros y carreras:

El estudiante 3

En la entrevista manifiesta que al construir una consulta escribe “(...) lo esencial, muchas veces, si es que la palabra (...)”, incluso explica cómo resolvió la tarea COVID 19. Sin embargo, en el experimento se observa que tanto para la tarea COVID 19 y Elecciones 2021, utilizó la pregunta como estrategia para elaborar sus consultas. En cambio, para la tarea “Noticias Falsas” su estrategia fue utilizar palabras clave. Lo anterior sugiere una oposición

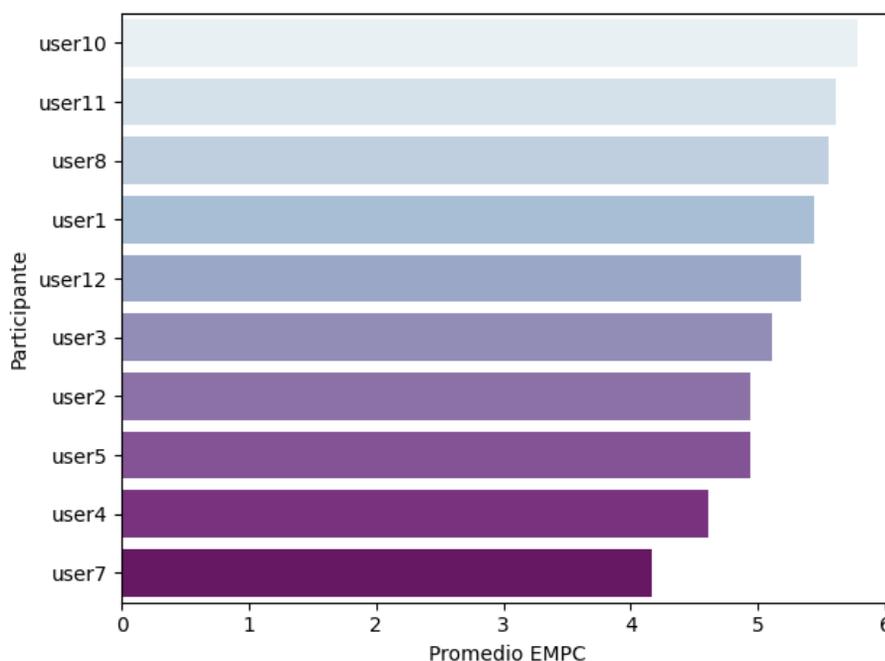
entre el discurso del estudiante y sus acciones concretas durante una búsqueda de información. Esto podría estar motivado por la rapidez inherente de la búsqueda de información que dificulta la reflexión del usuario. En cuanto a la selección de información en las tareas COVID 19 y Noticias Falsas el estudiante guarda fuentes de medios y un blog, pero para la tarea Elecciones 2021 no guarda páginas webs. No obstante, solo es posible observar el uso de información para la tarea Noticias Falsas. Las páginas webs usadas en esta tarea se ubican en el puesto 1 y 2 de la lista de resultados, lo que coincide con lo declarado en la entrevista. Lo anterior podría indicar que el estudiante considera que tiene un menor conocimiento del dominio en la tarea Noticias Falsas, sin embargo, recolecta información y la resume con sus propias palabras. Por el contrario, en las restantes tareas no es posible determinar qué fuentes utilizó para elaborar sus respuestas. Pero es interesante mencionar que para la tarea COVID 19 entregó un rango de cifras amplio que no se aproxima a la cifra oficial consignada en varias páginas webs disponibles en su lista de resultados. Vale decir, el estudiante dispuso de diversidad de fuentes con cifras correctas, sin embargo, responde con un rango de cifras errado. Lo anterior, podría considerarse como la operación del sesgo de Dunning-kruger (Dunning y Kruger, 1999). Este sesgo corresponde a la sobreestimación errónea de su conocimiento o capacidad en un área específica, producto de una falta de metacognición que le dificulta evaluar sus propias habilidades con precisión. No obstante, por ser un caso observado durante la triangulación de los datos del experimento y las entrevistas, no es posible confirmar la ocurrencia de dicho sesgo sin contrastarlo con la reflexión del estudiante respecto de sus decisiones en esta tarea específica.

El estudiante 5

Utiliza parte de la consigna de la tarea para elaborar sus consultas en las tareas COVID 19 y Noticias Falsas. Tal como menciona en la entrevista, en la tarea COVID 19 utiliza las comillas para forzar que en los resultados esté la palabra “muertes”. Esta estrategia fue eficiente, ya que en la primera consulta el algoritmo ubicó en el primer resultado la página web del Ministerio de Salud (MINSAL). De esta manera el estudiante accedió rápidamente a la cifra oficial de dicho ministerio, siendo la ruta esperada para resolver esta tarea. En cambio, en la tarea Elecciones 2021 usa la estrategia de pregunta, sin embargo, elabora un

resumen con sus propias palabras logrando una respuesta satisfactoria. Finalmente, en la tarea Noticias Falsas el estudiante, nuevamente, utiliza la primera fuente del listado de resultados, pero esta vez corta y pega las recomendaciones de la página web del medio de comunicación consultado. Este cambio podría estar asociado a un menor conocimiento del dominio o bien al cansancio, debido a que era la última tarea por resolver.

Ilustración 21: Resultados escala EMPC



La estudiante 10

La participante destaca con el mayor resultado en la escala EMPC (Ilustración 21). Lo que se observa también en la confianza mostrada durante la entrevista en sus habilidades de búsqueda. La estudiante considera que su amplio vocabulario y su competencia escrita facilitan la búsqueda y aprovechamiento de la información disponible. En esa línea manifestó que usualmente su estrategia de búsqueda es hacer una pregunta y destaca que lo hace como si se dirigiera a una persona. En efecto, utiliza esta estrategia en las tareas Elecciones 2021 y Noticias Falsas. En cambio, en la tarea COVID 19 elabora una consulta con términos propios que luego reformula incorporando nuevas palabras para delimitar la búsqueda. No obstante,

al igual que la mayoría de los participantes selecciona los primeros resultados propuestos por el algoritmo. Pero, para la última tarea elabora su resumen utilizando 3 páginas webs guardadas en su biblioteca. Esta participante tal como lo manifiesta en su entrevista demuestra mayor uso de vocabulario y habilidades de redacción.

La estudiante 11

Esta estudiante también destaca por una alta motivación a pensar críticamente (Ilustración 21). Durante la entrevista explica que suele elaborar consultas precisas “(...) porque si yo busco una pregunta muy amplia no busca lo que yo requiero (...)”, sin embargo, en todas sus consultas elabora frases o preguntas extensas que incluyen artículos y preposiciones. Al igual que la mayoría de los participantes elabora una única consulta y selecciona las fuentes que se ubican en las primeras posiciones de la lista de resultados. Respecto del uso de la información en la tarea COVID 19 es la única respuesta que utiliza la cifra de cálculo de muertes del medio digital “Ciper” que cuestiona la cifra oficial. Esto indicaría que la participante leyó el artículo completo, ya que las conclusiones del medio están en los últimos párrafos. Asimismo, en la tarea Elecciones 2021 resume con sus propias palabras la información. En cambio, en la tarea Noticias falsas corta y pega parte de la información de dos páginas webs consultadas. Si bien no elabora un resumen propio, al menos, utiliza información de dos fuentes. Este comportamiento podría estar condicionado por el orden temporal de la tarea y/o el dominio del tema.

Tabla 16: Trayectoria tarea "COVID 19"

	Tipo consulta	Consultas	Fuentes consultadas (Biblioteca)	Fuentes citadas	Ranking
1	resume consigna	Cantidad de fallecidos por covid19 en chile durante el año o	pauta.cl	Pauta	1
2	resume consigna/ aumenta datos	informe epidemiológico covid 19 chile/ informe epidemiológico covid 19 chile/ muertes por covid 19 chile 2020 ciper3	minsal.cl/ gob.cl	Ciper	1
3	Pregunta	Cuántas personas murieron en 2020 a causa del covid? /Cuántas personas murieron en 2020 a causa del covid en Chile? /Cuántas personas murieron hasta el 31 de diciembre 2020 a causa del covid en Chile?	bbc.com/ chile.as.com	n/a	n/a
4	palabras clave/aumenta datos	covid-19/covid-19 estadísticas nacionales/muertes en Chile el año 2020 por covid/numero de casos y muertes en Chile el año 2020	news.google/es.wikipedia.org/ine.es/diario concepción/chile.as.com/gob.cl	Minsal	2
5	resume consigna	"muertes" covid-19 covid 19 coronavirus/fallecimientos covid 19 Chile/fallecimientos covid 19 Chile 2020/fallecimientos covid 19 Chile	emol.com/minsal.cl	Minsal	1
7	resume consigna	fallecidos covid-19 Chile	minsal.cl	Minsal	1
8	resume consigna	cantidad de fallecidos por COVID-19 en Chile/cantidad de fallecidos por COVID-19 en Chile 2020	minsal.cl/24horas.cl	CNN	10
10	frase con tópicos propios	taza de mortalidad por COVID 19 en Chile/taza de mortalidad por COVID 19 en Chile el año 2020	24horas.cl/emol.com	CNN	1
11	resume consigna	cantidad de fallecidos por covid 19 en el año 2020	ciperchile	Ciper	1
12	resume consigna/ aumenta datos	fallecidos por covid-19 en Chile 2020/fallecidos por covid-19 en Chile en el 2020/fallecidos por covid-19 en Chile anual en el 2020/fallecidos por covid-19 en Chile durante todo el 2020	emol.com/24horas.cl	Ciper	2

Fuente: Elaboración propia

Tabla 17: Trayectoria tarea "Elecciones 2021"

	Tipo consulta	Consultas	Fuentes consultadas (Biblioteca)	Fuentes citadas	Ranking
1	resume consigna	Razones para el cambio de fecha de elecciones abril 2021	la cuarta.com	La Cuarta	3
2	palabras clave	Elecciones chile 2021	0	Canal Senado	1
3	Pregunta	Porque cambiaron la fecha de las elecciones abril 2021?	0	n/a	n/a
4	resume consigna	cambio de fecha en las elecciones 2021	rockandpop.cl/ diario concepción	Radio Rock and Pop	1
5	Pregunta	porque se cambio la fecha de elecciones 2021	pudahuel.cl	Radio Pudahuel	3
7	palabras clave	elecciones abril 2021 chile	elinformadorchile.cl	El informador Chile	10
8	palabras clave	Elecciones Abril 2021	0	CNN español	1
10	Pregunta	por que se cambiaron las fechas de elecciones abril 2021 en Chile?	radio.uchile.cl	Radio Universidad de Chile	2
11	Pregunta	porque se cambio de fechas las elecciones de abril 2021	elecciones2021.servel.cl	Servel	1
12	resume consigna	cambio de fechas elecciones abril 2021 chile	rockandpop.cl	Radio Rock and Pop	2

Fuente: Elaboración propia

Tabla 18: Trayectoria tarea "Noticias Falsas"

	Tipo consulta	Consultas	Fuentes consultadas (Biblioteca)	Fuentes citadas	Ranking
1	resume consigna y agrega información	Características de una noticia falsa, como reconocer noticias falsas	clasesdeperiodismo.com	Blog Clases de Periodismo	3
2	palabras clave propias	fake news	lavanguardia.com/bbc.com	lavanguardia/bbc	2 y 3
3	palabras clave	Características noticias falsas	clasesdeperiodismo.com	Blog Clases de Periodismo y BBC	1 y 2
5	resume consigna	características de noticias falsas	bbc.com	BBC	1
4	palabras clave propias	fake news	lavanguardia.com	La Vanguardia	2
7	palabras clave propias	características fake news	telecinco.es	Telecinco	2
8	resume consigna	Características de una noticia falsa	bbc.com/clasesdeperiodismo.com	Blog Clases de Periodismo	1
10	Pregunta con tópicos propios	que son las fake news?	unamglobal.unam.mx/telecinco.es/ ionos.es	n/a	2, 3 y 4
11	resume consigna	características de una noticia falsa	clasesdeperiodismo.com/brainly.lat	Brainly (sitio de pregunta/respuesta) y Blog Clases de Periodismo.	1 y 4
12	resume consigna	cuales son las características de las noticias falsas	clasesdeperiodismo.com/ bbc.com	Blog Clases de Periodismo	1

Fuente: Elaboración propia

6. Discusión

Para responder a las preguntas de investigación analizamos las trayectorias de búsqueda de los estudiantes en la plataforma GoNSA2 y las contrastamos con las respuestas de sus entrevistas. En la primera pregunta ¿Cuáles son las rutinas de búsqueda en internet que reconocen los estudiantes universitarios? Los participantes señalan que el motor de búsqueda es una herramienta cotidiana y que aprenden a usarla por sí mismos desde la niñez. Su uso es esencialmente académico y, por tanto, la frecuencia de uso está directamente relacionada con el número de tareas o dudas académicas. En este punto, manifestaron que su uso se intensificó durante la pandemia, debido a las dificultades ocasionadas por la educación remota. No obstante, durante las entrevistas emergieron usos no esperados por los investigadores, tales como, la búsqueda de información sobre intereses personales como el cine, videojuegos o descargas de programas computacionales. También, algunos estudiantes declararon usar la caja de búsqueda del motor para confirmar la ortografía de algunas palabras.

La segunda pregunta de investigación ¿Cuáles son las percepciones de los estudiantes universitarios respecto de los motores de búsqueda en internet? Durante las entrevistas se observó que los estudiantes utilizaban con mucha frecuencia esta herramienta y, en general, tenían una actitud positiva respecto de esta. Por este motivo, se indagó respecto de su valoración. Las respuestas fueron confirmatorias de la percepción de la entrevistadora, ya que los estudiantes expresaban que el buscador facilita el acceso a información relevante, destacando su rapidez. Es más, los estudiantes valoran también la selección del algoritmo, específicamente Google. Para los entrevistados, la selección del algoritmo generalmente cumple con lo esperado y, en los casos que declaran no recibir resultados satisfactorios, los atribuyen a su consulta. En esta idea, las estrategias para elaborar sus consultas responden a juicios intuitivos basados en la práctica. Explican que, normalmente, intentan realizar consultas breves porque según experiencia este tipo de consultas ofrece mejores resultados. Dicen no entender muy bien cómo funciona el algoritmo, pero suponen que más palabras amplían las posibilidades, lo que los conduce a información diversa que los confunde. Se aprecia que los estudiantes valoran que la búsqueda sea precisa y huyen de la información diversa. Incluso, varios detallan que buscan información breve y de fácil lectura. Si la página web es muy extensa, continúan la búsqueda.

En la tercera pregunta de investigación buscábamos identificar los posibles sesgos que operan durante la búsqueda de información. En general, es posible observar la presencia de sesgos o juicios intuitivos como los define Cortada (2018) debido a que los participantes no manifiestan un conocimiento razonado respecto de las estrategias de uso de los buscadores, sino que sus acciones se basan en su experiencia o en suposiciones. El principal hallazgo es la operación del sesgo de familiaridad y el sesgo de ranking del listado de resultados en todas las tareas, independiente del tipo de conocimiento y de los resultados de la escala EMPC. En la tarea COVID 19 los estudiantes reformulan sus consultas, pero al seleccionar un resultado eligen los que se ubican en las primeras posiciones (Tabla 15), lo que indicaría la operación del sesgo de ranking (Noble, 2018). Los participantes tienden a seleccionar las páginas web que se ubican en las primeras posiciones, evidenciando la confianza en la priorización del algoritmo de búsqueda. Este sesgo actúa en conjunto con el sesgo de familiaridad (Matute, 2019) con Google, lo que potencia una confianza sin reservas en la priorización del algoritmo, favoreciendo una actitud irreflexiva. Es más, la búsqueda de información los remite a una experiencia positiva “(...) lo he usado desde siempre y no lo he cambiado porque para mí funciona”, esto favorece que durante la búsqueda en internet los participantes bajen las alertas. Cabe mencionar que la plataforma GoNSA2 no permite visualizar el motor de búsqueda asociado, sin embargo, la mayoría de los estudiantes indicó que solo ha utilizado Google, lo que facilitaría que mantengan su conducta habitual de búsqueda. También se observa el sesgo de confirmación, varios estudiantes declaran que uno de sus criterios de parada es encontrar información que confirme lo que, previamente, esperan encontrar. Por lo tanto, si la información que logran localizar no se adscribe a lo que piensan respecto del tópico de búsqueda extienden la búsqueda. En esta misma instancia opera el sesgo de información (Baron, 2000), ya que la detención de la búsqueda está sujeta a la estimación subjetiva de la cantidad de información recopilada.

En nuestra última pregunta de investigación nos interesaba saber ¿Cuál es la relación de las percepciones de las rutinas de búsqueda con su motivación a pensar críticamente y su comportamiento observado entre los estudiantes universitarios? Basados en sus respuestas, observamos detalladamente su desempeño en la sesión experimental para comprobar la congruencia entre lo que los participantes creen hacer y lo que realmente hacen cuando buscan información en internet.

La tarea COVID 19 es la de mayor recordación, en las entrevistas varios participantes detallaron la estrategia que siguieron para encontrar la respuesta. Esto, podría estar ocasionado por que es la primera tarea que resuelven o bien por la dificultad de los participantes para encontrar la cifra solicitada, considerando que esta tarea es la que presentó un mayor número de consultas. Lo que mostró un resultado contrario a lo descrito en la literatura, ya que la mayor exploración de la web estaría relacionada con las tareas de tipo exploratorias (Rutter et al., 2019). En general, las tareas de tipo factual normalmente no requieren de una exploración profunda de la web, ya que las respuestas son fáciles de encontrar. Sin embargo, en la tarea “COVID 19” los 10 participantes realizaron un total de 25 consultas, con un promedio de 2.5 consultas. A pesar de ello, 4 participantes entregaron una respuesta alejada de la cifra oficial de personas fallecidas por COVID 19 en Chile durante el año 2020. Asimismo, a pesar del mayor número de consultas no se observan reformulaciones novedosas. En general, elaboraron consultas muy similares o realizaron pequeñas modificaciones intentando delimitar la búsqueda. Por ejemplo, una de las estrategias más eficientes fue agregar datos como: año, país o apuntar a un medio de comunicación en particular. Este comportamiento se opone a las intenciones declaradas en las entrevistas donde los estudiantes sostienen que elaboran consultas breves para encauzar la búsqueda de información.

En las siguientes tareas -de tipo exploratorias- se esperaba una mayor profundización de la web, representada por un mayor número de consultas y reformulaciones (Wildemuth y Freund, 2012), con una importante correlación con comportamientos menos eficientes y un mayor esfuerzo de la persona (Kalyani y Gadiraju, 2019). Sin embargo, en las tareas Elecciones 2021 y Noticias falsas, los estudiantes realizaron una única consulta. Otro aspecto interesante es que se observa que las estrategias de elaboración de consultas no cambian según el tipo de tarea, por lo que los estudiantes no muestran un mayor esfuerzo de búsqueda en tareas exploratorias. Podríamos asumir que la reducción de las consultas y/o el bajo esfuerzo podría estar influenciado por el orden temporal de la tarea. Sin embargo, sin la intención de desconocer el posible efecto del cansancio, se observa una mayor similitud de los resultados (o menor diversidad) en la primera consulta lo que consideramos tiene un efecto observable (medible) sobre la decisión de detener la búsqueda de información. Sin embargo, en ambas tareas también observamos que el nivel de exploración en la web es

dependiente de la mediatización de la tarea. En definitiva, en ambos casos, con una mayor influencia en la tarea “Noticias falsas”, el efecto de la mediatización del tópico de la búsqueda facilita la rápida localización de la información solicitada, debido a la mayor heterogeneidad de los sitios web de la lista de resultados. Específicamente, en la tarea Noticias Falsas, los estudiantes tienen acceso a varios medios de comunicación internacionales con recomendaciones a sus lectores para reconocer las *fake news*.

Lo anterior, influye directamente en la conformidad de los estudiantes con los resultados alcanzados indicado en el bajo número de consultas (Tabla 16 y 17). No obstante, al relacionar el comportamiento durante la búsqueda con las entrevistas se observa que, efectivamente, los estudiantes indican que suelen hacer pocas consultas y elegir las páginas webs mejor posicionadas en la lista de resultados. Por lo que la tarea COVID 19 sería una búsqueda atípica que requirió reformulaciones debido a la dificultad de encontrar la cifra correspondiente al rango de tiempo solicitado en la tarea.

Como caso atípico el estudiante 7, que obtuvo el resultado más bajo de la escala EMPC, para la tarea Elecciones 2021 selecciona una página web ubicada en la décima posición del listado de resultados. Lo anterior, podría sugerir que su menor conocimiento del dominio de la búsqueda lo impulsó a profundizar en su búsqueda o bien el medio de comunicación es conocido por el participante. En cualquier caso, el estudiante ajusta su estrategia de selección para resolver a conformidad la tarea propuesta. Este comportamiento sugiere que su percepción respecto de su motivación al pensamiento crítico es más alta que la reportada en sus respuestas. Por el contrario, en estudiantes que registran resultados más altos se observa un comportamiento con tendencia a un bajo esfuerzo en la resolución de las tareas de búsqueda de información. En general, se observan diferencias entre la percepción de los estudiantes respecto de su motivación a pensar críticamente y su comportamiento de búsqueda, lo que podría estar motivado por un bajo autoconcepto, propio de la edad de los participantes.

7. Conclusiones

Las percepciones y el uso de los motores de búsqueda están basadas en su experiencia práctica y del razonamiento intuitivo de los estudiantes, por lo que se observa el uso de heurísticas para resolver sus búsquedas de información. Es importante destacar que los participantes no mencionan en sus respuestas el término “sesgos” o similares, lo que podría indicar desconocimiento de su ocurrencia en la web. No obstante, en el análisis de las entrevistas se observa la operación de cuatro sesgos cognitivos. Primero, el sesgo de familiaridad con Google y con las páginas webs conocidas o utilizadas con anterioridad. Segundo, el sesgo de ranking, el cual se vincula estrechamente con la confianza reportada por la decisión del algoritmo de Google, dado que la mayoría solo conoce y ha utilizado este buscador tienden a repetir el patrón de comportamiento de búsqueda durante el experimento, aun cuando la plataforma no se vincula con Google. Tercero, el sesgo de información (Baron, 2000), entendido como el volumen de datos y ejemplos posibles de recopilar. Cuarto, el sesgo de confirmación, varios de los entrevistados declaran que buscan información que confirme sus presupuestos respecto del tema de búsqueda.

En general, la mayoría de los participantes admite dificultades para construir consultas efectivas y destaca la complejidad de seleccionar información creíble. No obstante, la idea de que “(...) Básicamente, todo lo puedes encontrar ahí.” está presente en la mayoría de las entrevistas. Los estudiantes asumen mediante un juicio intuitivo que la indexación de Google es casi ilimitada y que contiene toda la información posible de acumular. Por lo tanto, solo es necesario hacer la consulta precisa para alcanzar el mejor resultado posible. Al parecer, el sesgo de familiaridad con Google tiende a intensificar la confianza en la selección del algoritmo, por lo que no cuestionan sus decisiones. Es más, consideran que si no obtienen resultados óptimos es debido a la deficiencia de sus consultas, lo que atribuyen principalmente a su insuficiente vocabulario.

No obstante, los estudiantes son conscientes de que en la web es posible encontrar información falsa, con errores o poco fiable. Pero, los criterios para determinar si una información es creíble son escasos y fáciles de superar. De las respuestas se desprende que para persuadir a los usuarios basta con difundir información escrita correctamente, con una fecha actualizada, usar referencias, utilizar un dominio que no sea sospechoso y evitar las

ventanas emergentes. Si la información supera estas barreras, posiblemente, sea aceptada como creíble.

Otro hallazgo interesante para considerar es que los participantes no perciben la existencia de sesgos en la web, al menos no lo manifiestan durante las entrevistas. Es más, los estudiantes estiman a los algoritmos como entidades neutras que seleccionan el mejor resultado de manera rápida y eficiente. Estos resultados nos alientan a continuar en esta línea de trabajo, promoviendo la reflexión de las personas como usuarias habituales de plataformas de búsqueda de información.

8. Limitaciones

La muestra corresponde a estudiantes de primer año de universidad, por lo que no fue posible observar diferencias significativas entre los participantes de las distintas disciplinas. Por lo anterior, consideramos que sería apropiado replicar este estudio en cursos superiores para observar si la disciplina de estudio de los participantes influye en la percepción, sesgos y la elaboración de estrategias durante la búsqueda de información. Asimismo, sería interesante examinar si la formación académica actúa como facilitadora de procesos de reflexión en los estudiantes. Para ello es de nuestro interés elaborar un estudio longitudinal que permita establecer una correlación entre la formación académica y la reflexión crítica de los estudiantes durante la búsqueda de información en internet.

CONCLUSIONES

1. Conclusiones

En el primer capítulo de este estudio se problematiza respecto de las implicancias de los algoritmos de búsqueda como configuradores de subjetividades, a partir del perfilado algorítmico de usuarios de estas plataformas. En la literatura constatamos que se han definido diversos alcances de la personalización en la vida y comportamiento de las personas (Boyd et al., 2014; Chander, 2016; Parisier, 2017; Siles et al., 2019). Sin embargo, sin interés de demonizar estas tecnologías, es innegable que no son neutras y tienen efectos sobre el comportamiento y la vida cotidiana de sus usuarios, aunque sean difíciles de observar. En distintas investigaciones analizadas en esta tesis se ha establecido que la información disponible en internet no es accesible en igual medida para todos y que depende de distintas variables establecidas para facilitar el funcionamiento técnico de los buscadores. Si a esto le sumamos las brechas técnicas y de uso de los propios usuarios, identificadas también en esta investigación, se desmitifica el acceso universal a la información prometido en los inicios de internet, pero que parece permanecer inalterable en el imaginario colectivo.

De esta discusión surge la necesidad de un acercamiento empírico al objeto de estudio a través del diseño de dos instancias de experimentación que permitieran observar el comportamiento de estudiantes universitarios chilenos durante sesiones de búsqueda de información.

El primer trabajo empírico mostró que es posible alcanzar, mediante una búsqueda de información, un mayor número fuentes de edición validada si el tópico de la tarea está presente en la agenda mediática. No obstante, el listado de resultados SERP integra un alto número de páginas webs sin procesos editoriales, el que aumenta en temas que no están en los medios de comunicación. Dichos resultados sugieren que la agenda mediática influiría sobre el número de consultas realizadas durante una búsqueda de información. En este primer cuasi experimento se observó que las listas de resultados mostraron un ranking casi intacto con mínimas fluctuaciones entre los participantes. Es decir, los estudiantes accedieron casi a las mismas fuentes con diferentes consultas, evidenciando la acción de los filtros del algoritmo, tal como se planteó en el primer capítulo de esta tesis. No obstante, un caso atípico

que realizó 13 consultas para una misma tarea demostró que un mayor número de consultas incide en la posibilidad de alcanzar mayor variedad de fuentes.

Considerando lo anterior, en el segundo acercamiento empírico se elaboraron tareas de búsqueda con temáticas presentes en la agenda mediática nacional, para favorecer la posibilidad de encontrar resultados asociados a la tarea. Lo anterior, nos permitió comprobar que la mediatización del contexto de la tarea aumenta considerablemente las fuentes disponibles en la web y reduce la emisión de consultas y/o reformulaciones. Además, el segundo estudio empírico revela que tal como se describe en la literatura los estudiantes son permeables a los sesgos introducidos por la lista de resultados del algoritmo (Wu & Kelly, 2015). En ambos procesos experimentales los estudiantes tendieron a confiar en las decisiones del algoritmo del motor de búsqueda. Esta confianza en la selección algorítmica se evidencia en:

El sesgo de disponibilidad reduce el número de consultas por tarea. En general, los estudiantes reducen el número de consultas cuando se sienten conformes con los resultados alcanzados. Este sesgo opera en tareas donde existen muchas fuentes con información. En el segundo estudio empírico se mostró que independiente del tipo de tarea, los estudiantes aumentan el número de consultas cuando las fuentes de información son diversas. Los participantes disminuyeron sus consultas en las tareas exploratorias “Elecciones abril 2021” y “Noticias Falsas”. En cambio, aumentaron la exploración en la tarea “COVID 19” que solicitaba una cifra específica. En el caso de esta tarea se solicitaba la cifra de fallecidos por COVID 19 en el año 2020, sin embargo, la diversidad de fuentes disponibles y la controversia mediática en los modos de cálculo de los fallecimientos dificultó la búsqueda. Es necesario considerar que el tópico de esta tarea tenía una alta presencia en la agenda mediática, por lo que los participantes podían tener conocimiento de las fuentes de información válidas. De este modo, aquellos estudiantes que acudieron directamente a la fuente oficial alcanzaron un resultado en su primera búsqueda. En este caso constatamos que el conocimiento previo permitió que realizarán una consulta eficiente (Rutter et al., 2019), influyendo sobre la decisión de detener la búsqueda de manera temprana. Sin embargo, en los casos que no conocían una fuente válida, la búsqueda se extendió incrementando la interacción de la persona con el motor de búsqueda, lo que implicaría un mayor esfuerzo en el proceso (Liu, Liu & Belkin, 2013; Chi, 2019). Sin embargo, no podemos constatar si la búsqueda se

extendió por falta de conocimiento del dominio o bien por falta de credibilidad en las fuentes, lo cual implicaría un conocimiento parcial del dominio, pero no de las fuentes de información oficial, lo que podría explicarse en la mediatización del tema COVID 19. Asimismo, en las entrevistas los estudiantes declaran que si conocen la fuente de información acceden a ella directamente desde el buscador y detienen la búsqueda en una única consulta.

La selección de los primeros resultados provistos por el algoritmo. En este estudio se observó que la mayoría de los estudiantes tendieron a seleccionar las páginas webs ubicadas en las primeras posiciones. Sin embargo, se constata que en la tarea “COVID 19” la selección se extiende hacia resultados ubicados más abajo en la lista, posiblemente por los motivos expuestos en el párrafo anterior, pero siempre en la primera página de resultados. El sesgo de ranking se observó también en las entrevistas donde la mayoría de los participantes manifiestan, abiertamente, usar los primeros resultados porque confían en la selección del algoritmo (Wu & Kelly, 2015; Schultheiß et al., 2018; Su et al., 2018; Noble, 2018; Liu y Zhang, 2019). Aunque declaran no conocer los criterios que usa el algoritmo para elaborar el listado de resultados, se sienten conformes con su desempeño, basados únicamente en su experiencia personal. A nuestro pesar, se demuestra uno de los principales planteamientos del primer capítulo, las personas creemos que mediante una búsqueda en internet podemos acceder a toda la información disponible sobre un tema. Debido, principalmente, al desconocimiento del funcionamiento técnico y de las estrategias de posicionamiento de las páginas web.

También se observa y caracteriza el sesgo de anclaje en cuatro instancias del proceso de búsqueda de información: En la elaboración de consultas, en la selección de fuentes de información, en la construcción de la respuesta a la tarea solicitada y en el sesgo de algoritmo por la consulta del usuario. En este sesgo las variables saturación y motivación al pensamiento crítico (EMPC) observadas de manera independiente no son significativas. Sino que el sesgo de anclaje tiende a incrementar en función del tópico de la tarea. Sin embargo, al analizar la interacción de ambas variables (saturación y EMPC) se observa que influyen de manera significativa sobre la operación del sesgo de anclaje durante la construcción de la respuesta a la tarea.

En la selección de las fuentes se observa un mayor nivel de sesgo de anclaje sobre todo en la relación consulta-título, ya que los estudiantes seleccionan las fuentes, basándose principalmente en el título del resultado. Sin embargo, las diferencias de anclaje dependen principalmente del menor conocimiento previo del contexto de la tarea, lo que aumentaría el sesgo de anclaje con el título del resultado.

En el análisis de las entrevistas se confirma la operación de sesgos durante el proceso de búsqueda de información. Sin embargo, en las respuestas los participantes no se mencionan la palabra “sesgos” o similares, lo que podría indicar que no son conscientes de los sesgos en sus decisiones mientras buscan información en la web. No obstante, es posible observar la operación cuatro sesgos cognitivos. Primero, se reitera el sesgo de ranking. Los estudiantes explican que la selección de páginas webs depende directamente del orden de la lista de resultados. En general, manifiestan que seleccionan los primeros resultados porque confían en la selección del buscador. Muchos de ellos hacen referencia al algoritmo como la entidad que selecciona los mejores resultados, los más actuales y usados por otros usuarios. Lo anterior, les da garantía de confianza sobre la fuente. Pero, también declaran fijarse en el título y el *snippet* de los primeros resultados para confirmar la cercanía del resultado con su consulta. Segundo, el sesgo de familiaridad. Este sesgo es el que presenta la más alta adhesión por parte de los participantes. Todos los estudiantes declaran usar Google como buscador favorito desde siempre. Algunos conocen otros buscadores, pero no los consideran válidos. Al indagar respecto de la predilección por Google las respuestas remiten a emociones y experiencias que los vinculan al buscador, incluso manifiestan que su uso está ligado al acostumbramiento. Aunque, esta relación de cercanía está vinculada a la integración de Google con Chrome, tal como lo evidencian en sus respuestas. Tercero, se advierte el sesgo de información en la decisión de detener la búsqueda. Los participantes declaran que uno de los criterios más relevantes es alcanzar una cantidad suficiente de información, lo que incluye el contenido y ejemplos que les permitan “sentir” que tienen la cantidad de información suficiente. El problema es que esta suficiencia no está vinculada a la calidad ni relevancia sino a la cantidad de información. Vale decir, los participantes tienden a creer que más información les permite resolver de mejor forma una tarea. Cuarto, se advierte la ocurrencia del sesgo de confirmación. Los estudiantes explican que al buscar información confían en las fuentes que les proporcionan la información que ellos esperan obtener, es decir, aquella que

confirme lo que ellos esperan encontrar, por lo que tienden a aumentar el número de consultas si la información no responde a lo esperado. Aunque, explican que, normalmente, Google responde de manera acertada a sus consultas, por lo que rara vez no encuentran la información que buscan, lo que aumenta su satisfacción con el motor de búsqueda. Los participantes entienden que el algoritmo ha personalizado sus intereses y responde a su comportamiento general, sin embargo, en sus respuestas muestran conformidad, incluso, complacencia con la personalización de sus intereses. Efectivamente, entre los participantes no se observan cuestionamientos a las decisiones algorítmicas. Tan solo en un par de estudiantes se rescatan ciertas prácticas de resistencia, a partir del entrenamiento del algoritmo para búsquedas en inglés. Estos estudiantes son conscientes de que la personalización implica la reducción del espacio de búsqueda, por lo que usan el inglés como estrategia de resistencia. Incluso, uno de ellos declara que el tipo de información que busca con frecuencia se encuentra más actualizada en dicho idioma.

En general, los participantes no evidencian habilidades técnicas de acceso y localización eficientes de la información durante sus búsquedas. Es más, en las entrevistas la mayoría de los participantes calificó sus competencias de búsqueda de información como “bajas” o “promedio” y, las respuestas de aquellos que respondieron tener “buenas” habilidades no fueron enfáticas sino más bien dubitativas. Además, los estudiantes manifiestan que han aprendido a buscar en casa, mediante el uso cotidiano desde edades tempranas. Por lo que sus habilidades se han construido a partir de su propia experiencia sin una guía pedagógica.

Finalmente, se concluye que en los estudiantes universitarios que participaron de esta investigación la operación de sesgos cognitivos ocurre en la interacción de la motivación al pensamiento crítico con el nivel de saturación de información. Sin embargo, se observó que la variable emergente tipo de tarea, presenta una mayor influencia sobre el proceso de búsqueda. Asimismo, durante la etapa cualitativa se detectó que el sesgo de familiaridad con Google es determinante en las decisiones de búsqueda de los estudiantes. Lo anterior, puede ser visto como una apatía o pereza mental que condiciona nuestra relación con la tecnología. Se infiere que los participantes de este estudio tienden a manifestar una actitud pasiva que delega en la tecnología decisiones que anteriormente tomaban los seres humanos.

2. Implicancias

En los trabajos previos, los acercamientos metodológicos del estudio de los motores de búsqueda involucran la manipulación de los resultados de búsqueda. Por ejemplo, se diseñan condiciones de laboratorio para el desarrollo de los experimentos con los participantes, lo que limita la comprensión de las complejidades de la búsqueda de información. En cambio, la propuesta metodológica de esta tesis favorece el estudio del comportamiento de búsqueda en un ambiente naturalista. Donde los participantes realizan sus búsquedas, desde sus propios dispositivos, a través de la interfaz de búsqueda GoNSA2. Esto permite que los estudiantes reciban la lista de resultados que selecciona el algoritmo sin manipulación ni intervención de la investigadora. Asimismo, la automatización de los procesos de la aleatorización y el almacenamiento de las trazas de comportamiento permite que los estudiantes busquen sin sentirse observados. Lo que permite obtener las estrategias de búsqueda habituales de los estudiantes en una aproximación empírica que asemeja las condiciones naturales de búsqueda.

En cuanto a las implicancias a nivel metodológico. Este estudio propone un diseño metodológico transdisciplinar para el análisis del consumo de información como una práctica de aprendizaje autónomo. Por lo que este diseño metodológico es transferible a estudios de prácticas de consumo de distintos tipos de información, a través del análisis de las decisiones de los usuarios frente a condiciones tecnológicas y ambientales.

A nivel práctico, las estrategias de consumo de información por parte de los participantes nos entregan indicios acerca de ilimitadas estrategias que utilizan para incluir/excluir páginas web. En este sentido, en las entrevistas constatamos que en los participantes no existe una actitud crítica respecto del desempeño de los algoritmos, sino que la crítica apunta a sus habilidades de búsqueda. Es más, reconocen que el uso de los motores de búsqueda, en concreto Google, es un hábito difícil de reemplazar. Esta familiaridad con Google sesga la posibilidad de buscar información en otros buscadores u otros medios. Por ello, es necesario que las instituciones de educación superior promuevan el uso crítico de la información y, además, enseñen a sus estudiantes cómo acceder a buscadores especializados.

En esta tesis se logra caracterizar la operación de sesgos durante todo el proceso de búsqueda. Desde la emisión de las consultas al motor de búsqueda hasta la elaboración de las respuestas. En definitiva, los estudiantes se muestran permeables a los sesgos durante la búsqueda, lo que pone en riesgo la autonomía en el aprendizaje. Por ejemplo, se ha evidenciado que la mediatización de la tarea influye sobre el esfuerzo de los participantes en la búsqueda de información. Es decir, mientras más presencia de la agenda mediática tenga la tarea, mayor es la confianza de los estudiantes sobre los resultados. Asimismo, la operación de sesgos cognitivos en todas las tareas de búsqueda, interactúan con la motivación a pensar críticamente y el nivel de saturación de los resultados. Por lo que esta caracterización de las prácticas de resolución de tareas en entornos digitales nos permite personalizar un modelo de formación de competencias mediáticas e informacionales con un claro enfoque en el pensamiento crítico y en el rol de los medios de comunicación como fuente de información. De esta manera, se plantea un avance en la formación de las competencias digitales e informacionales de los estudiantes de educación superior.

3. Líneas de trabajo futuras

Examinar la influencia de programas de formación de alfabetización informacional sobre el consumo de medios durante la búsqueda de información. En este estudio se han evidenciado las dificultades del proceso de búsqueda de información, así como las consecuencias que tiene sobre las decisiones de los estudiantes respecto del esfuerzo invertido en la búsqueda y la emergencia de los sesgos cognitivos, principalmente, en la selección de información para la elaboración de una respuesta. Asimismo, se ha evidenciado una alta presencia de la agenda mediática en los resultados de búsqueda y, asimismo, una tendencia a seleccionar medios de información como fuentes válidas. Por lo tanto, una primera línea de trabajo futuro es el estudio de la relación entre los programas de alfabetización mediática e informacional y las prácticas de consumo de medios de información.

Analizar la presencia de la agenda mediática durante las prácticas de consumo de información en la web. En este trabajo doctoral se puso en tensión la influencia de los algoritmos sobre la dieta informativa de los usuarios. En los estudios cuasi experimentales

se evidenció una alta presencia de medios de comunicación en las listas de resultados propuestas por el motor de búsqueda. Por lo tanto, una segunda línea de trabajo se sustenta en la necesidad de conocer la influencia de la agenda mediática sobre las decisiones en las prácticas de consumo de los estudiantes y de los algoritmos de búsqueda en la selección de medios de comunicación como fuentes de información.

Estudiar la credibilidad de los medios por parte de los estudiantes durante la búsqueda de información. En este estudio se obtuvo evidencia de la relación entre las características mediáticas de las tareas propuestas y la duración de las sesiones experimentales. Dependiendo de la mediatización de la tarea los participantes aumentaban o reducían el esfuerzo de búsqueda. Por lo que es necesario profundizar respecto de las razones detrás del aumento del esfuerzo, y establecer si se vincula al interés con el tópico o bien sobre la credibilidad sobre las fuentes de información. En consecuencia, se vislumbra una tercera línea de trabajo respecto del estudio de la credibilidad de los medios de comunicación a través de las prácticas de búsqueda de información en tareas altamente mediatizadas.

REFERENCIAS

- Abad, G. L. (2012). *Exclusión digital y límites de la comunicación mediada*. Trípodos, (31), 111-123.
- Aguaded, I., & Romero-Rodríguez, L. M. (2015). Mediamorfosis y desinformación en la infoesfera: Alfabetización mediática, digital e informacional ante los cambios de hábitos de consumo informativo= Mediamorphosis and misinformation in the infosphere: media, digital and information literacy face of changes in information consumption habits. *EKS* 16 (1), 44-57. <http://dx.doi.org/10.14201/eks20151614457>
- Alexander, S., Kernohan, G., & McCullagh, P. (2004). 3.3. Self Directed and Lifelong Learning. Hovenga, E. J., & Mantas, J. (Eds.), *Global Health Informatics Education* (Vol. 109, pp. 152-166). IOS Press.
- American Library Association. (2003). The nine information literacy standards for student learning. <http://home.mca.k12.pa.us/~kozlowskiw/library/The%20Nine%20Information%20Literacy%20Standards%20for%20Student%20Learn%E2%80%A6.pdf>
- Aparici R. (2011). Introducción: Educomunicación más allá del 2.0. Aparici, R. (Ed.). *Educomunicación: más allá del 2.0*. (1, pp.1-15). Gedisa.
- Arguello, J., Wu, W. C., Kelly, D., & Edwards, A. (2012, August). Task complexity, vertical display and user interaction in aggregated search. *In Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval* (435-444).
- Baeza-Yates, R. (2018). Bias on the web. *Communications of the ACM*, 61(6), 54-61. <https://doi.org/10.1145/3209581>
- Baeza-Yates, R. (2020). Bias in search and recommender systems. *In Fourteenth ACM Conference on Recommender Systems* (2-2). <https://doi.org/10.1145/3383313.3418435>
- Balabantaray, R. C (2017). Evaluation of web search engines based on ranking of results and its features. *International Journal of Human Computer Interaction*, 10(4), 392-405. <https://doi.org/10.1504/IJICT.2017.084337>
- Banerjee, M., Zlatkin-Troitschanskaia, O., & Roeper, J. (2020). Narratives and their impact on students' information seeking and critical online reasoning in higher education

- economics and medicine. In *Frontiers in Education*, 5, p. 570625.
<https://doi.org/10.3389/feduc.2020.570625>
- Barilá S. (2003). Regulation of TV Contents in Argentina. Cecilia von Feilitzen C. and Carlsson U. (Eds.), *Promote or Protect? Perspectives on Media Literacy and Media Regulations* (pp. 213-220). The International Clearinghouse on Children, Youth and Media NORDICOM Göteborg University. ISBN 91-89471-23-7
- Baron, J. (2000). *Thinking and deciding* (3ª edición). Cambridge University Press.
- Bawden, D. and Robinson, L. (2009). The dark side of information: overload, anxiety and other paradoxes and pathologies. *Journal of information science*, 35(2), 180-191.
<https://doi.org/10.1177/0165551508095781>
- Bazarova, N. N., & Choi, Y. H. (2014). Self-disclosure in social media: Extending the functional approach to disclosure motivations and characteristics on social network sites. *Journal of Communication*, 64(4), 635-657.
<https://doi.org/10.1111/jcom.12106>
- Bhaskar, M. (2018). *Curaduría: el poder de la selección en un mundo de excesos*. Fondo de Cultura Económica.
- Black, B. (2012). An overview of a programme of research to support the assessment of Critical Thinking. *Thinking Skills and Creativity*, 7(2), 122-133.
<https://doi.org/10.1016/j.tsc.2012.04.003>
- Boyd, D., Levy, K., & Marwick, A. (octubre 2014). The networked nature of algorithmic discrimination. *Data and Discrimination: Collected Essays*. Open Technology Institute. <http://www.danah.org/papers/2014/DataDiscrimination.pdf>
- Brandão, G. D. S., Santos, J. O., & Borges, J. (2020). Comportamento infocomunicacional: um diagnóstico entre estudantes para a promoção de competências infocomunicacionais. *Informação & informação*. Londrina, PR. 25(4), 670-696.
<https://doi.org/10.5433/1981-8920.2020v25n4p670>
- Breakstone, J., Smith, M., Connors, P., Ortega, T., Kerr, D., & Wineburg, S. (2021). Lateral reading: College students learn to critically evaluate internet sources in an online course. *The Harvard Kennedy School Misinformation Review* 2, (1), 1-17.
<https://doi.org/10.37016/mr-2020-56>

- Buckingham, D. (2007). Media education goes digital: an introduction. *Learning, Media and technology*, 32(2), 111-119. <https://doi.org/10.1080/17439880701343006>
- Buckingham, D. (2008). *Más allá de la tecnología. Aprendizaje infantil en la era de la cultura digital*. Manantial.
- Buckingham D. & Domaille K. (2003). Where Are We Going and How Can We Get There? General Findings from the UNESCO Youth Media Education Survey 2001. Cecilia von Feilitzen C. and Carlsson U. (Eds.), *Promote or Protect? Perspectives on Media Literacy and Media Regulations* (pp. 41-54). The International Clearinghouse on Children, Youth and Media NORDICOM Göteborg University. ISBN 91-89471-23-7
- Burbules, N. C. (2012). El aprendizaje ubicuo y el futuro de la enseñanza. *Encounters on education*, 13 (212), 3-14. <https://doi.org/10.24908/eoe-ese-rse.v13i0.4472>
- Bucher, T. (2018). *If... then: Algorithmic power and politics*. Oxford University Press.
- Burrell, J. (2016). How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1). <https://doi.org/10.1177/2053951715622512>
- Cachorro, G. (2008). Cuerpo y subjetividad: Rasgos, configuraciones y proyecciones. En *Jornadas de Cuerpo y Cultura de la UNLP*. Universidad Nacional de La Plata. Facultad de Humanidades y Ciencias de la Educación. Departamento de Educación Física. http://www.memoria.fahce.unlp.edu.ar/trab_eventos/ev.697/ev.697.pdf
[Consulta: 10 agosto 2019]
- Canclini, N. G. (2004). *Diferentes, desiguales y desconectados: mapas de la interculturalidad*. Gedisa.
- Cano-Orón, L. (2019). Dr. Google, what can you tell me about homeopathy? Comparative study of the top10 websites in the United States, United Kingdom, France, Mexico and Spain. *Profesional de la Información*, 28 (2) 1-12. <https://doi.org/10.3145/epi.2019.mar.13>
- Caplan, R., & Boyd, D. (2016). Who controls the public sphere in an era of algorithms. *Mediation, Automation, Power*, 1-19. https://datasociety.net/wp-content/uploads/2016/05/MediationAutomationPower_2016-1.pdf

- Carrasco, R. F. (mayo de 2003). La cara oculta de Internet. *Hipertext. net*, (1) <http://goo.gl/G6DMkh>
- Castells, M. (2009). *Comunicación y Poder*. Alianza.
- Castells, M. (2003). La Era de la Información. *Vol. I: La sociedad Red*. Alianza Editorial.
- Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. CA: Sage Publications.
- Cobo, C. y Moravec, J. (2011). *Aprendizaje Invisible. Hacia una nueva ecología de la educación*. Publicacions Edicions de la Universitat de Barcelona.
- Cope, W. & Kalantzis, M. (2010). *Ubiquitous learning*. University of Illinois Press.
- Cornellá, A. (2004). *Infoxicación: Buscando un orden en la información*. Infonomía.
- Cormier, D. (2008). Rhizomatic education: Community as curriculum. *Innovate: Journal of online education*, 4(5). file:///C:/Users/instalacion2020_1/Downloads/article_104239.pdf
- Cortada de Kohan, N. (2008). Los sesgos cognitivos en la toma de decisiones. *International Journal of Psychological Research* 2008, 1, (1), 68-73. <https://doi.org/10.21500/20112084.968>
- Coslado, Á. B. (2012). Educomunicación: desarrollo, enfoques y desafíos en un mundo interconectado. *Foro de educación*, 10(14), 157-175.
- Chander, A. (2016). The racist algorithm. *Mich. L. Rev.*, 115, 1023. <http://repository.law.umich.edu/mlr/vol115/iss6/13>.
- Chen, H. J., Lin, H. J., Wu, M. C., Tang, H. J., Su, B. A., & Lai, C. C. (2021). The implementation of an active surveillance integrating information technology and drive-through coronavirus testing station for suspected COVID-19 cases. *Journal of Infection*, 82(2), 282-327. <https://doi.org/10.1016/j.jinf.2020.08.051>
- Chi, Y. (2019). Examining and Supporting Laypeople's Learning in Online Health Information Seeking. In *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval* (pp. 425-428). <https://doi.org/10.1145/3295750.3298975>
- Clark D. (2003). A Practical Response to Classification of Convergent Media in the Australian Context. The Combined Guidelines for Films and Computer Games. Cecilia von Feilitzen C. and Carlsson U. (Eds.), *Promote or Protect? Perspectives on Media Literacy*

- and Media Regulations* (pp. 231-236). The International Clearinghouse on Children, Youth and Media NORDICOM Göteborg University. ISBN 91-89471-23-7
- Danaher, J., Hogan, M. J., Noone, C., Kennedy, R., Behan, A., De Paor, A., ... & Shankar, K. (2017). Algorithmic governance: Developing a research agenda through the power of collective intelligence. *Big data & society*, 4(2), 2053951717726554. <https://doi.org/10.1177/2053951717726554>
- Daniel, M.F. (2002). *La philosophie et les enfants*. Belin.
- de Kunder, M. (2019). The size of the World Wide Web (The Internet). <https://www.worldwidewebsite.com/>
- da Silva Almeida, L. and Rodrigues Franco, A. H. (2011). Critical thinking: Its relevance for education in a shifting society, 29 (1), 175-195. <http://hdl.handle.net/20.500.12799/2356>
- Davenport T. H. & Beck. J. C. (2001). The Attention economy. *Ubiquity* 2001, May, Article 6 (May 1 - May 31, 2001). <https://doi.org/10.1145/375348.376626>
- Dias, P. (2014). From ‘infoxication’ to ‘infosaturation’: a theoretical overview of the cognitive and social effects of digital immersion. *Ambitos*, 24. <http://hdl.handle.net/10400.14/14939>
- Deleuze, G., & Guattari, F. (1987). *A thousand plateaus: capitalism and schizophrenia*. University of Minnesota Press.
- Delors, J. (1995). *La educación encierra un tesoro. Informe sobre la educación para el siglo XXI*. <https://bibliotecadigital.mineduc.cl/handle/20.500.12365/10446?show=full>
- Díaz-Larenas, C. H., Ossa-Cornejo, C. J., Palma-Luengo, M. R., Martín, L. S., Gromiria, N., & Boudon Aranedá, J. I. (2019). El concepto de pensamiento crítico según estudiantes chilenos de pedagogía. *Sophia, colección de Filosofía de la Educación*, (27), 275-296. <https://doi.org/10.17163/soph.n27.2019.09>
- Domínguez, P. S., & López, Á. J. G. (2019). Data activism versus algorithmic control. New governance models, old asymmetries. *IC Revista Científica de Información y Comunicación*, (16), 183-208. <https://dx.doi.org/10.12795/IC.2019.i01.06>
- Druin, A., Foss, E., Hatley, L., Golub, E., Guha, M. L., Fails, J., & Hutchinson, H. (2009). How children search the internet with keyword interfaces. In *Proceedings of the 8th*

- International conference on interaction design and children* (89-96).
<https://doi.org/10.1145/1551788.1551804>
- Duarte Torres, S., Hiemstra, D., Weber, I., & Serdyukov, P. (2012). Query recommendation for children. In *Proceedings of the 21st ACM international conference on Information and knowledge management* (2010-2014). <https://doi.org/10.1145/2396761.2398562>
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology*, 77(6), 1121–1134. <https://doi.org/10.1037/0022-3514.77.6.1121>
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual review of psychology*, 53(1), 109-132.
- Eisenberg, M. B., y Berkowitz, R. E. (2000). *Teaching Information & Technology Skills: The Big6 [TM] in Secondary Schools*. Linworth Publishing, Inc.
- Elder, L. and Paul, R. (1994). *Critical thinking development: A stage theory*. Foundation for Critical Thinking.
- Ennis, R. H., Millman, J. and Tomko, T. N. (1985). *Cornell critical thinking tests level X & level Z: Manual*. Pacific Grove. Midwest Publications.
- Ennis, R. H. (2002). An outline of goals for a critical thinking curriculum and its assessment. <https://www.Faculty.Ed.Uiue.Edulrhennis>.
- Ennis, R. H. (2011). The nature of critical thinking: An outline of critical thinking dispositions and abilities. *University of Illinois*, 2(4), 1-8.
- Facione, P. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. The Delphi Report.
- Facione, P. A. (2000). The disposition toward critical thinking: Its character, measurement, and relationship to critical thinking skill. *Informal logic*, 20(1).
- Facione, P. (2007). Pensamiento Crítico: ¿Qué es y por qué es importante? *Insight assessment*, 23 (1), 22-56.
- Fernández-Vicente, A. (2020). Hacia una teoría crítica de la razón algorítmica. *Palabra Clave*, 23(2), 1-26. <https://doi.org/10.5294/pacla.2020.23.2.2>
- Ferrés, J. (2014). *Las pantallas y el cerebro emocional*. Colección Comunicación Educativa. Editorial Gedisa.

- Ferrés, J., & Piscitelli, A. (2012). La competencia mediática: propuesta articulada de dimensiones e indicadores. *Comunicar*, 38(19), 75-82.
- Foo, S., Majid, S., Mokhtar, I. A., Zhang, X., Chang, Y. K., Luyt, B., & Theng, Y. L. (2014). Information literacy skills of secondary school students in Singapore. *Aslib Journal of Information Management*, 66 (1), 54–76. <https://doi.org/10.1108/AJIM-08-2012-0066>
- Fornas, R. (2003). La cara oculta de Internet [en línea]. "Hipertext.net",1. <http://www.hipertext.net>
- Foss, E. and Druin, A. (2014). Children's Internet Search: Using Roles to Understand Children's Search Behavior. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 6(2), 1–106. <https://doi.org/10.2200/S00591ED1V01Y201408ICR034>
- Flick, U. (2007). *Introducción a la investigación cualitativa*. Morata.
- Gaitán, J. A., & Piñuel, J. L. (1997). *Técnicas de investigación social. Elaboración y registro de datos*. Síntesis.
- García-Martín, S. and Cantón-Mayo, I. (2019). Use of technologies and academic performance in adolescent students/Usó de tecnologías y rendimiento académico en estudiantes adolescentes. *Comunicar*, 27 (59), 73-81. <https://doi.org/10.3916/C59-2019-07>
- Gehl, R. W. (2014). Power/freedom on the dark web: A digital ethnography of the Dark Web Social Network. *New Media & Society*, 18 (7), 1219-1235. <https://doi.org/10.1177/1461444814554900>
- Ghenai, A., Smucker, M. D., & Clarke, C. L. (2020, March). A think-aloud study to understand factors affecting online health search. In *Proceedings of the 2020 conference on human information interaction and retrieval* (273-282).
- Goldhaber, M. H. (1997). The attention economy and the net. *First Monday*. 2, 4-7. <https://journals.uic.edu/ojs/index.php/fm/article/download/519/440>
- González-Osorio, G. (2021). Competencias digitales de alumnos de primer semestre de la UPV. *Kinesis Revista Veracruzana de Investigación Docente*, 6 (6), 92-111. <https://www.revistakinesis.com/index.php/journal/article/view/98>

- Gozálvez-Pérez, V., & Aguaded-Gómez, J. I. (2012). Educación para la autonomía en sociedades mediáticas. *Anàlisi: quaderns de comunicació i cultura*, (45), 0001-14. <https://doi.org/10.7238/a.v0i45.1326>
- Gonzalo C. (2015). Posicionamiento web y dinámicas de información en motores de búsqueda: propuestas de análisis y estudio comparativo de visibilidad de contenidos digitales en el caso de procesos electorales. *Hipertext.net [Online]*, 13. <http://raco.cat/index.php/Hipertext/article/view/294133/389445>
<https://doi.org/10.2436/20.8050.01.28>
- Gutiérrez-Coba, L., Coba-Gutiérrez, P., & Gómez-Díaz, J. A. (2020). Las noticias falsas y desinformación sobre el Covid-19: análisis comparativo de seis países iberoamericanos. *Revista Latina de Comunicación Social*, (78), 237-264. <https://doi.org/10.4185/RLCS-2020-1476>
- Haim, M., Arendt, F. and Scherr, S. (2017). Abyss or shelter? On the relevance of web search engines' search results when people Google for suicide. *Health communication*, 32 (2), 253-258. <https://doi.org/10.1080/10410236.2015.1113484>
- Haim, M., Graefe, A. and Brosius, H. B. (2018). Burst of the filter bubble? Effects of personalization on the diversity of Google News. *Digital journalism*, 6 (3), 330-343. <https://doi.org/10.1080/21670811.2017.1338145>
- Halavais, A. (2017). *Search engine society*. John Wiley & Sons.
- Halpern, D. F. (2006). The nature and nurture of critical thinking. En RJ Sternberg, R. Roediger y DF Halpern (Eds.), *Critical thinking in psychology* (pp. 1-14).
- Han, B. C. (2013). *La Sociedad de la Transparencia*. Herder Editorial.
- Han, B. C. (2014a). *En el Enjambre*. Herder Editorial.
- Han, B. C. (2014b). *Psicopolítica*. Herder Editorial.
- Ho, H. J., Lim, W. Y., Ang, B., & Chow, A. (2021). Use of surveillance technology to enhance exposure management for healthcare workers during the COVID-19 pandemic. *The Journal of Hospital Infection*, 107, 101. <https://doi.org/10.1016%2Fj.jhin.2020.09.024>
- Jenkins, Henry (2006). *Convergence culture. Where old and new media collide*. New York University Press. ISBN: 978 0 814742952 <https://goo.gl/AvrMUA>

- Jenkins, H., Ford, S., & Green, J. (2015). *Cultura transmedia: la creación de contenido y valor en una cultura en red* (Vol. 60). Editorial Gedisa.
- Jiang, M. (2014a). Search concentration, bias, and parochialism: A comparative study of Google, Baidu, and Jike's search results from China. *Journal of Communication*, 64 (6), 1088-1110. <https://doi.org/10.1111/jcom.12126>
- Jiang, M. (2014b). The business and politics of search engines: A comparative study of Baidu and Google's search results of Internet events in China. *New Media & Society*, 16b (2), 212-233. <https://doi.org/10.1177/1461444813481196>
- Just, N., & Latzer, M. (2017). Governance by algorithms: reality construction by algorithmic selection on the Internet. *Media, Culture & Society*, 39b (2), 238-258. <https://doi.org/10.1177/0163443716643157>
- Kahneman, D. (2012). *Pensar rápido, pensar despacio*. Debate.
- Kalyani, R., & Gadiraju, U. (2019). Understanding user search behavior across varying cognitive levels. In *Proceedings of the 30th ACM conference on hypertext and social media* (123-132). <https://doi.org/10.1145/3342220.3343643>
- Kaplún, M. (1998). *Una pedagogía de la comunicación*. Ediciones de la Torre.
- Karppi, T., & Crawford, K. (2016). Social media, financial algorithms and the hack crash. *Theory, culture & society*, 33(1), 73-92. <https://doi.org/10.1177/0263276415583139>
- Kearns, M., & Roth, A. (2019). *The ethical algorithm: The science of socially aware algorithm design*. Oxford University Press.
- Kelly, D., & Azzopardi, L. (2015). How many results per page? A study of SERP size, search behavior and user experience. In *Proceedings of the 38th international ACM SIGIR conference on research and development in information retrieval* (183-192). <https://doi.org/10.1145/2766462.2767732>
- Kestin, T., van den Belt, M., Denby, L., Ross, K., Thwaites, J., & Hawkes, M. (2017). *Getting started with the SDGs in universities: a guide for universities, higher education institutions, and the academic sector*. <https://apo.org.au/sites/default/files/resource-files/2017-08/apo-nid105606.pdf>
- Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, communication & society*, 20(1), 14-29. <https://doi.org/10.1080/1369118X.2016.1154087>

- Kroll, J. A. (2015). *Accountable algorithms* (Doctoral dissertation, Princeton University).
<https://oar.princeton.edu/bitstream/88435/pr1r24q/1/AccountableAlgorithms.pdf>
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology*, 77(6), 1121. <https://doi.org/10.1037/0022-3514.77.6.1121>
- Kules, B. and Shneiderman, B. (2008). Users can change their web search tactics: Design guidelines for categorized overviews. *Information Processing & Management*, 44 (2), 463-484. <https://doi.org/10.1016/j.ipm.2007.07.014>
- Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American society for information science*, 42 (5), 361-371.
[https://doi.org/10.1002/\(SICI\)1097-4571\(199106\)42:5<361::AID-ASI6>3.0.CO;2-%23](https://doi.org/10.1002/(SICI)1097-4571(199106)42:5<361::AID-ASI6>3.0.CO;2-%23)
- Kuhn, D., & Weinstock, M. (2002). *What is epistemological thinking and why does it matter?* Lawrence Erlbaum Associates Publishers.
- Lewandowski D. (2017) Is Google Responsible for Providing Fair and Unbiased Results?. In: Taddeo M., Floridi L. (eds) *The Responsibilities of Online Service Providers. Law, Governance and Technology Series*, 31. Springer, Cham.
https://doi.org/10.1007/978-3-319-47852-4_4
- Li, X., & Dong, H. (2020). An Oligopoly Two-Stage-Game Model for Investigating the Search Engine Market. *International Journal of Performability Engineering*, 16(11).
<https://doi.org/10.23940/ijpe.20.11.p10.17811792>
- Liu, J., & Zhang, X. (2019). The role of domain knowledge in document selection from search results. *Journal of the Association for Information Science and Technology*, 70(11), 1236-1247. <https://doi.org/10.1002/asi.24199>
- Liu, J. (2019). A Reference-Dependent Model of Search Evaluation. In *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval* (405-408).
<https://doi.org/10.1145/3295750.3298970>

- Liu, J., Liu, C., & Belkin, N. (2013). Examining the effects of task topic familiarity on searchers' behaviors in different task types. In *Proceedings of the American Society for Information Science and Technology* 50, 1–10. American Society for Information Science. <https://doi.org/10.1002/meet.14505001033>
- Lodder, A. R., & Loui, R. P. (2018). Data algorithms and privacy in surveillance: On stages, numbers and the human factor. In *Research handbook on the law of artificial intelligence*. Edward Elgar Publishing. <https://doi.org/10.4337/9781786439055.00025>
- Marjanovic, O., & Cecez-Kecmanovic, D. (2017). Exploring the tension between transparency and datification effects of open government IS through the lens of Complex Adaptive Systems. *The Journal of Strategic Information Systems*, 26 (3), 210-232. <https://doi.org/10.1016/j.jsis.2017.07.001>
- Masterman L. (1993). *La enseñanza de los medios de comunicación*. Ediciones de la Torre.
- Manovich, L. (2014). *El software toma el mando*. Editorial UOC.
- Marchionini, G. (2006). Exploratory search: from finding to understanding. *Communications of the ACM*, 49 (4), 41-46. <https://doi.org/10.1145/1121949.1121979>
- Mattelart, A. (2007). *Historia de la sociedad de la información*. Paidós.
- Matute, H. (2019). *Nuestra mente nos engaña. Sesgos y errores cognitivos que todos cometemos*. Shackleton books.
- McBride, S. (1980). *Informe McBride. Voces múltiples, un solo mundo. Comunicación e información en nuestro tiempo*. Fondo de cultura económica.
- McPeck, J. E. (2016). *Teaching Critical Thinking. Dialogue and Dialectic*. Routledge. <https://doi.org/10.4324/9781315526492>
- Mesquita-Romero, W. A., Fernández-Morante, C., & Cebreiro-López, B. (2022). Alfabetización mediática crítica para mejorar la competencia del alumnado. *Comunicar*, (70), 47-57. <https://doi.org/10.3916/C70-2022-04>
- Morsy, Z. (Coord.) (1984). *La educación en materia de comunicación*. UNESCO.
- Möller, J., Trilling, D., Helberger, N., and van Es, B. (2018). Do not blame it on the algorithm: an empirical assessment of multiple recommender systems and their impact on content diversity. *Information, Communication & Society*, 21 (7), 959-977. <https://doi.org/10.1080/1369118X.2018.1444076>

- Nagel, M. T., Schäfer, S., Zlatkin-Troitschanskaia, O., Schemer, C., Maurer, M., Molerov, D., ... & Brückner, S. (2020). How do university students' web search behavior, website characteristics, and the interaction of both influence students' critical online reasoning? In *Frontiers in Education*, 5, p. 565062. <https://doi.org/10.3389/feduc.2020.565062>
- Ni Loideain, N. N., & Adams, R. (2020). From Alexa to Siri and the GDPR: the gendering of virtual personal assistants and the role of data protection impact assessments. *Computer Law & Security Review*, 36, 105366. <https://doi.org/10.1016/j.clsr.2019.105366>
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.
- Novin, A., & Meyers, E. (2017). Making sense of conflicting science information: Exploring bias in the search engine result page. In *Proceedings of the 2017 conference on conference human information interaction and retrieval* (175-184). <https://doi.org/10.1145/3020165.3020185>
- Organisation for Economic Co-operation and Development (2018). *Education at a Glance 2018: OECD Indicators*. OECD Publishing. <https://doi.org/10.1787/eag-2018-en>.
- Olivares-Rodríguez C. and Guenaga M. (2015) Detection of the Student Creative Behavior Based on Diversity of Queries. In: García-Chamizo J., Fortino G., Ochoa S. (eds) *Ubiquitous Computing and Ambient Intelligence. Sensing, Processing, and Using Environmental Information. UCAmI 2015. Lecture Notes in Computer Science, 9454*. Springer, Cham. https://doi.org/10.1007/978-3-319-26401-1_18
- Olivares-Rodríguez, C., Guenaga, M., Garaizar, P. (2017). "Automatic assessment of creativity in heuristic problem-solving based on query diversity". *Dyna*, 92, (4), 449-455. <http://dx.doi.org/10.6036/8243>
- Olivares-Rodríguez, C., Guenaga, M. and Garaizar, P. (2018). Using children's search patterns to predict the quality of their creative problem solving. *Aslib Journal of Information Management*, 70 (5), 538-550. <https://doi.org/10.1108/AJIM-05-2018-0103>

- Olivares-Rodríguez, C., Valdés-León, G., Vidal-Sepúlveda, M. (2021). “Influencia de la agenda mediática en las búsquedas de información: una experiencia con estudiantes universitarios”. *ReHuSo*, 6 (2),17-28. <https://doi.org/10.33936/rehuso.v6i2.3423>
- Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). *The PageRank citation ranking: Bringing order to the web* (Technical Report). Stanford InfoLab. <http://ilpubs.stanford.edu:8090/422/>>
- Palmiotto, F. (2021). The black box on trial: the impact of algorithmic opacity on fair trial rights in criminal proceedings. In *Algorithmic Governance and Governance of Algorithms* (49-70). Springer, Cham. https://doi.org/10.1007/978-3-030-50559-2_3
- Pariser, E. (2017). *El filtro burbuja: Cómo la web decide lo que leemos y lo que pensamos*. Taurus.
- Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.
- Pasquinelli, M. (2016). *The spike: On the growth and form of pattern police*. Nervous Systems.
- Pereira, S., Fillol, J., Moura, P. (2019). El aprendizaje de los jóvenes con medios digitales fuera de la escuela: De lo informal a lo formal. *Comunicar*, 27, (58), 41-50. <https://doi.org/10.3916/C58-2019-04>
- Piergiorgio, C. (2007). *Metodología y técnicas de investigación social*. McGraw-Hill.
- Potter, W. J. (2013). Review of literature on media literacy. *Sociology Compass*, 7(6), 417-435.
- Pothirattanachaikul, S., Yamamoto, T., Yamamoto, Y., & Yoshikawa, M. (2019). Analyzing the Effects of Document's Opinion and Credibility on Search Behaviors and Belief Dynamics. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management* (1653-1662). <https://doi.org/10.1145/3357384.3357886>
- Qureshi, A., Bokhari, S. A. H., Pirvani, M., & Dawani, N. (2015). Understanding and practice of evidence based search strategy among postgraduate dental students: a preliminary study. *Journal of Evidence Based Dental Practice*, 15 (2), 44-49. <https://doi.org/10.1016/j.jebdp.2014.08.002>

- Rasmussen, E. M. (2003). Indexing and retrieval for the Web. *Annual Review of Information Science and Technology (ARIST)*, 37, 91-124.
- Reimers, F. & Chung, C. (ed.) (2016). *Enseñanza y aprendizaje en el siglo XXI: metas, políticas educativas y currículo en seis países*. Fondo de Cultura Económica.
- Rieder, B. and Sire, G. (2014). Conflicts of interest and incentives to bias: A microeconomic critique of Google's tangled position on the Web. *New media & society*, 16 (2), 195-211. <https://doi.org/10.1177/1461444813481195>
- Rieder, B. (2016). *RankFlow*. <http://labs.polsys.net/tools/rankflow>
- Rivas Navarro, M. (2008). *Procesos cognitivos y aprendizaje significativo*. Comunidad de Madrid. Consejería de Educación. Viceconsejería de Organización Educativa.
- Roberge, J. (2018). What are algorithmic cultures? En R. Seyfert, & J. Roberge (Eds.), *Algorithmic cultures. Essays on meaning, performance and new technologies* (pp. 13-37). Routledge.
- Romero-Rodríguez, L. M. & Mancinas R. (2016). Sobresaturación informativa: visibilizar el mensaje institucional en tiempos de infoxicación en Luis Miguel Romero-Rodríguez (Ed), *Comunicación Institucional y Cambio Social. Claves para la comprensión de los factores relacionales de la comunicación estratégica y el nuevo ecosistema comunicacional* (pp. 111-137). Egregius.
- Romero-Rodríguez, L. M., & Aguaded, I. (2016). Consumo informativo y competencias digitales de estudiantes de periodismo de Colombia, Perú y Venezuela. *Convergencia*, 23 (70), 35-57.
- Rosman, T., Mayer, A. K., & Krampen, G. (2015). Combining self-assessments and achievement tests in information literacy assessment: empirical results and recommendations for practice. *Assessment & Evaluation in Higher Education*, 40 (5), 740-754. <https://doi.org/10.1080/02602938.2014.950554>
- Rowlands, I., Nicholas, D., Williams, P., Huntington, P., Fieldhouse, M., Gunter, B., ... and Tenopir, C. (2008). The Google generation: the information behaviour of the researcher of the future. *In Aslib proceedings* 60 (4), 290-310. <https://doi.org/10.1108/00012530810887953>
- Rutter, S., Blinzler, V., Ye, C., Wilson, M. L., & Twidale, M. D. (2019). Search tactics used in solving everyday how-to technical tasks: Repertoire, selection and tenacity.

- Information Processing & Management*, 56 (3), 919-938.
<https://doi.org/10.1016/j.ipm.2019.02.008>
- Sadikov, E., Madhavan, J., Wang, L., & Halevy, A. (2010). Clustering query refinements by user intent. In *Proceedings of the 19th international conference on World wide web* (841-850). <https://doi.org/10.1145/1772690.1772776>
- Sales, D., Cuevas-Cerveró, A., Gómez-Hernández, J. (2020). “Perspectives on the information and digital competence of Social Sciences students and faculty before and during lockdown due to Covid-19”. *Profesional de la información*, 29, (4),1-21.
<https://doi.org/10.3145/epi.2020.jul.23>
- Sautu, R., Boniolo, P., Dalle, P., y Elbert, R. (2005). *Manual de metodología: construcción del marco teórico, formulación de los objetivos y elección de la metodología*. CLACSO.
- Santos Jr, E., Nguyen, H., Zhao, Q., & Wang, H. (2003). User modelling for intent prediction in information analysis. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 47, (8), 1034-1038.
<https://doi.org/10.1177/154193120304700818>
- Saiz, C., & Rivas, S. (2008). Intervenir para transferir en pensamiento crítico. *Praxis*, 10 (13), 129-149.
- Saiz, C., Rivas, S. F. and Olivares, S. (2015). Collaborative Learning Supported by Rubrics Improves Critical Thinking. *Journal of the Scholarship of Teaching and Learning*, 15 (1), 10-19. <https://doi.org/10.14434/josotl.v15i1.12905>
- Domínguez, P. S., & López, Á. J. G. (2019). Data activism versus algorithmic control. New governance models, old asymmetries. *IC Revista Científica de Información y Comunicación*, (16), 183-208. <https://dx.doi.org/10.12795/IC.2019.i01.05>
- Schultheiß, S., Sünkler, S., & Lewandowski, D. (2018). We still trust in Google, but less than 10 years ago: an eye-tracking study. *Information Research: An International Electronic Journal*, 23 (3), n3. <http://informationr.net/ir/23-3/paper799.html>
- Scolari, C. A. (2016). Alfabetismo transmedia: estrategias de aprendizaje informal y competencias mediáticas en la nueva ecología de la comunicación. *Telos: Revista de pensamiento sobre Comunicación, Tecnología y Sociedad*. 193, 13-23.

- Scriven, M. & Paul, R. (1992, November). Critical thinking defined. In *Handout given at Critical Thinking Conference*, Atlanta, GA.
- Scriven, M. and Paul, R. (2003). *Defining Critical Thinking: A draft statement prepared for the National Council for Excellence in Critical Thinking Instruction*. University/defining. html.
- Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big data & society*, 4(2). <https://doi.org/10.1177/2053951717738104>
- Sefton-Green, J. (2006). *Literature Review in Informal Learning with Technology Outside School Media* (Report 7). Nesta Futurelab Series. https://www.researchgate.net/publication/32231568_Literature_Review_in_Informal_Learning_with_Technology_Outside_School
- Sefton-Green, J. (2013). *Learning at Not-School: A Review of Study, Theory, and Advocacy for Education in Non-Formal Settings*. The John D. and Catherine T. MacArthur Foundation. Reports on Digital Media and Learning. MIT Press (BK).
- Şendurur, E., & Yildirim, Z. (2015). Students' web search strategies with different task types: an eye-tracking study. *International Journal of Human-Computer Interaction*, 31 (2), 101-111. <https://doi.org/10.1080/10447318.2014.959105>
- Serrano-Puche, J. (2013). Una propuesta de dieta digital: repensando el consumo mediático en la era de la hiperconectividad. *Fonseca, Journal of Communication*, 7 (2), 156-175. <https://dadun.unav.edu/bitstream/10171/35984/1/Una%20propuesta%20de%20dieta%20digital-FJC.pdf>
- Serrano-Puche, J. (2014). Por una dieta digital, hábitos mediáticos saludables contra la “obesidad informativa”. *Ámbitos. Revista Internacional de Comunicación*, (24). <http://dx.doi.org/10.12795/Ambitos.2014.i24.08>
- Shenk, D. (2003). Information overload, concept of. E. Science. *Encyclopedia of International Media and Communications*, 2, 395-405.
- Shipard S. (2003). A Brief Look at the Regulation of the Broadcast Media in Australia. Cecilia von Feilitzen C. and Carlsson U. (Eds.), *Promote or Protect? Perspectives on Media Literacy and Media Regulations* (pp. 237-242). The International Clearinghouse on Children, Youth and Media NORDICOM Göteborg University. ISBN 91-89471-23-7

- Siemens, G. (2006). *Knowing knowledge*. Lulu. com.
- Sibilia, P. (2012). *El hombre postorgánico: cuerpo, subjetividad y tecnologías digitales*. Fondo de cultura económica.
- Sibilia, P. (2012). *La intimidad como espectáculo*. Fondo de cultura económica.
- Siegel, H. (2010) Critical thinking, en Peterson, P.; Baker, E. y McGaw, B. (eds.) *International Encyclopedia of Education*. Elsevier (pp. 141-145).
- Siles, I., Espinoza-Rojas, J., Naranjo, A., & Tristán, M. F. (2019). The mutual domestication of users and algorithmic recommendations on Netflix. *Communication, Culture & Critique*, 12 (4), 499-518. <https://doi.org/10.1093/ccc/tcz025>
- Simon, H. A. (1957). *Models of man; social and rational*. Wiley.
- Speier, C., Valacich, J. S., & Vessey, I. (1999). The influence of task interruption on individual decision making: An information overload perspective. *Decision sciences*, 30 (2), 337-360. <https://doi.org/10.1111/j.1540-5915.1999.tb01613.x>
- Standish, P., & Thoilliez, B. (2018). El pensamiento crítico en crisis. Una reconsideración pedagógica en tres movimientos. Teoría De La Educación. *Revista Interuniversitaria*, 30(2), 7–22. <https://doi.org/10.14201/teoredu302722>
- Steiner, M., Magin, M., Stark, B., & Geiß, S. (2020). Seek and you shall find? A content analysis on the diversity of five search engines' results on political queries. *Information, Communication & Society*, 25(2), 217-241. <https://doi.org/10.1080/1369118X.2020.1776367>
- Su, N., He, J., Liu, Y., Zhang, M., & Ma, S. (2018, February). User intent, behaviour, and perceived satisfaction in product search. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining* (547-555). <https://doi.org/10.1145/3159652.3159714>
- Suárez-Gonzalo, S. (2017). Big social data: límites del modelo notice and choice para la protección de la privacidad. *El Profesional de la Información*. 26 (2), 283-92. <https://doi.org/10.3145/epi.2017.mar.15>
- Tello, L. (2013). Intimidad y extimidad en las redes sociales: las demarcaciones éticas de Facebook= Intimacy and Extimacy in Social Networks: Ethical Boundaries of Facebook, *Comunicar*, 21 (41), 205-213. <http://dx.doi.org/10.3916/C41-2013-20>
- Tung, C. A. and Chang, S. Y. (2009). Developing critical thinking through literature reading. Feng Chia. *Journal of Humanities and Social Sciences*, 19 (3), 287-317.

- Tversky, A. & Kahneman, D. (1983). "Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment". *Psychological review*, 90 (4), 293-315. <https://doi.org/10.1037/0033-295X.90.4.293>
- Urgo, K., Arguello, J., & Capra, R. (2019). Anderson and krathwohl's two-dimensional taxonomy applied to task creation and learning assessment. In *Proceedings of the 2019 ACM SIGIR International Conference on Theory of Information Retrieval* (117-124). <https://doi.org/10.1145/3341981.3344226>
- Usta, A., Altıngövdü, I., y Vidinli, I. (2014). How k-12 students search for learning?: analysis of an educational search engine log. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval* (pp. 1151–1154). <https://doi.org/10.1145/2600428.2609532>
- Vidal-Sepúlveda, M., & Olivares-Rodríguez, C. (2021). ¿Quién elige mi dieta informativa? La datificación automatizada de personas en el ambiente digital. *Información, cultura y sociedad*, (44), 143-154. <https://dx.doi.org/10.34096/ics.i44.8627>
- Valenzuela, J. & Nieto, A. M. (2008). Motivación y Pensamiento Crítico: Aportes para el estudio de esta relación. *REME*, 9, (28). <http://reme.uji.es/articulos/numero28/article3/article3.pdf>
- Valenzuela, J., Muñoz Valenzuela, C., Silva-Peña, I., Gómez Nocetti, V., & Precht Gandarillas, A. (2015). Motivación escolar: Claves para la formación motivacional de futuros docentes. *Estudios pedagógicos*, 41 (1), 351-361. <http://dx.doi.org/10.4067/S0718-07052015000100021>
- Varis, T., & Pérez Tornero, J. M. (2012). *Alfabetización mediática y nuevo humanismo*. UOC.
- Vásquez, A. (2016). Byung-Chul Han: la sociedad de la transparencia, psicopolítica y autoexplotación neoliberal: de lo viral-inmunológico a lo neuronal-estresante. En *Reflexiones Marginales*. No. 31. <https://reflexionesmarginales.com.mx/blog/2016/01/29/byung-chul-han-2/>
- Vilches, L. (Ed.). (2011). *La investigación en comunicación: métodos y técnicas en la era digital*. Gedisa.
- von Feilitzen, C. (2004). Promote or Protect? Perspectives on Media Literacy and Media Regulations: Introduction. Cecilia von Feilitzen C. and Carlsson U. (Eds.),

- Promote or Protect? Perspectives on Media Literacy and Media Regulations* (pp. 9-21). The International Clearinghouse on Children, Youth and Media NORDICOM Göteborg University. ISBN 91-89471-23-7
- Webber, W., Moffat, A., & Zobel, J. (2010). A similarity measure for indefinite rankings. *ACM Transactions on Information Systems (TOIS)*, 28(4), 1-38.
- Wildemuth, B. M. and Freund, L. (2012, October). Assigning search tasks designed to elicit exploratory search behaviors. In *Proceedings of the Symposium on Human-Computer Interaction and Information Retrieval* (p. 4). ACM. <https://doi.org/10.1145/2391224.2391228>
- Wilson, C., Grizzle, A., Tuazon, R., Akyempong, K., & Cheung, C. K. (2011). *Alfabetización Mediática e informacional: Curriculum para profesores*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000216099>
- Wu, W. C., Kelly, D., Edwards, A. and Arguello, J. (2012, August). Grannies, tanning beds, tattoos and NASCAR: Evaluation of search tasks with varying levels of cognitive complexity. In *Proceedings of the 4th Information Interaction in Context Symposium* (254-257). ACM. <https://doi.org/10.1145/2362724.2362768>
- Wu, W. C., & Kelly, D. (2015). Online search stopping behaviors: An investigation of query abandonment and task stopping. *Proceedings of the American Society for Information Science and Technology*, 51 (1), 1-10. <https://doi.org/10.1002/meet.2014.14505101030>
- Xing, W., & Ghorbani, A. (2004, May). Weighted pagerank algorithm. In *Proceedings. Second Annual Conference on Communication Networks and Services Research*. (305-314). IEEE. <https://doi.org/10.1109/DNSR.2004.1344743>
- Yagci, N., Sünkler, S., Häußler, H., & Lewandowski, D. (2022). A Comparison of Source Distribution and Result Overlap in Web Search Engines. arXiv preprint arXiv:2207.07330. <https://doi.org/10.48550/arXiv.2207.073>
- Yang, Y. T. C. (2012). Cultivating critical thinkers: Exploring transfer of learning from pre-service teacher training to classroom practice. *Teaching and Teacher Education*, 28 (8), 1116-1130. <https://doi.org/10.1016/j.tate.2012.06.007>

- Yamamoto, Y., Yamamoto, T., Ohshima, H., & Kawakami, H. (2018, May). Web access literacy scale to evaluate how critically users can browse and search for web information. In *Proceedings of the 10th ACM Conference on Web Science* (97-106). <https://dl.acm.org/doi/10.1145/3201064.3201072>
- Yoganarasimhan, H. (2020). Search personalization using machine learning. *Management Science*, 66(3), 1045-1070. <https://doi.org/10.1287/mnsc.2018.3255>
- Zach, L. (2005). When is 'enough' enough? Modeling the information-seeking and stopping behavior of senior arts administrators. *Journal of the American Society for Information*, 56(1), 23–35. <https://doi.org/10.1002/asi.20092>