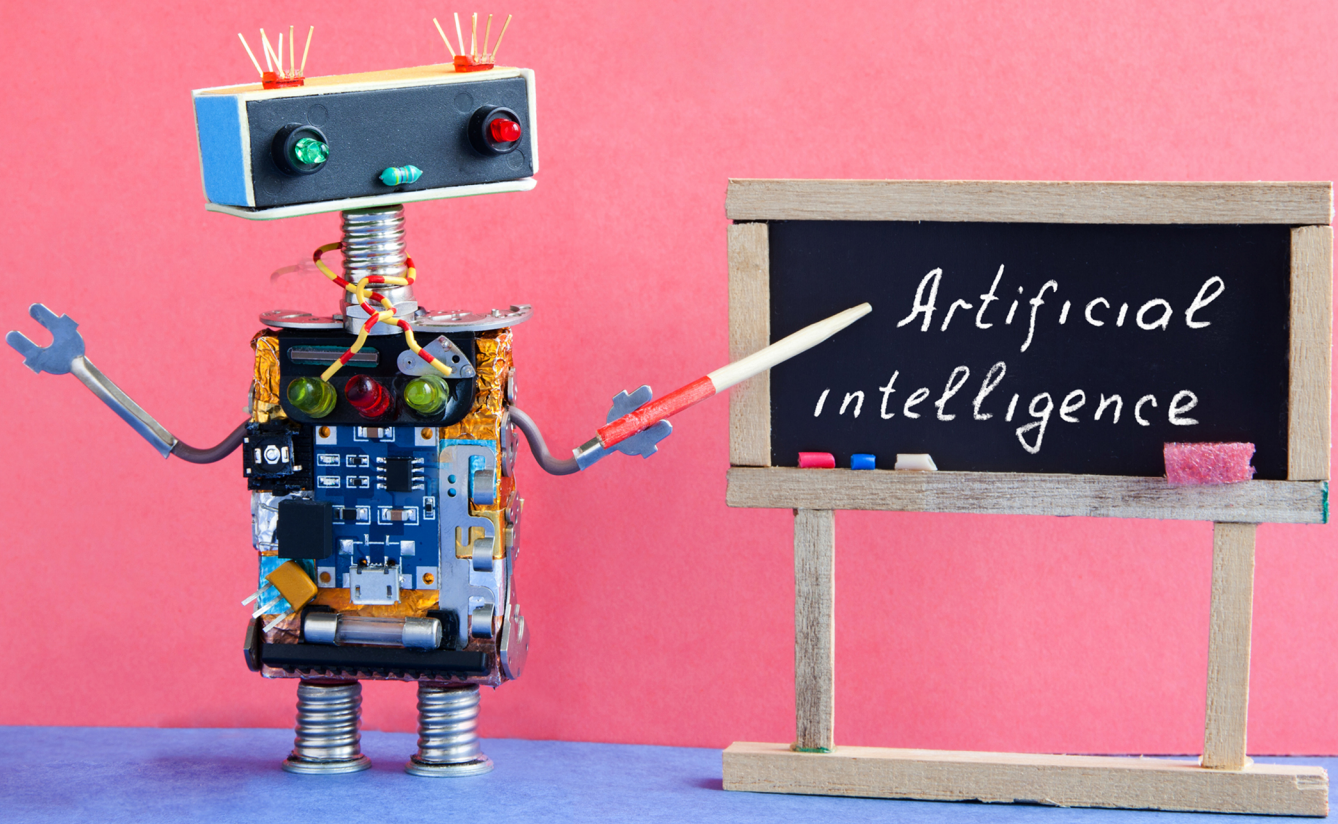
Comunicar
Media Education Research Journal, n. 74, vol. XXXI
www.comunicarjournal.com

Education for digital citizenship: Algorithms, automation and communication



English Edition

© COMUNICAR, 74, XXXI

MEDIA EDUCATION RESEARCH JOURNAL
REVISTA CIENTÍFICA DE COMUNICACIÓN Y EDUCACIÓN

ISSN: 1134-3478 / DL: H-189-93 / e-ISSN: 1988-3293
n. 74, vol. XXXI (2023-1), 1st Quarter, January, 1st, 2023

INDEXED INTERNATIONAL SCIENTIFIC JOURNAL

https://www.revistacomunicar.com/index.php?contenido=factor-de-impacto&idioma=en

JOURNAL CITATION REPORTS (JCR)

JCR 2021 (2022-23): Q1. JIF: 5.725. JCI: 2.94. 5 Year Impact Factor: 5.715. Inmediacy Index: 2,000; Eigenfactor Score: 0.00252. Article Influence Score: 1.330; Journal Impact Factor (JIF): Education: Q1 (position 18 from 267; 1st Spanish and Iberoamerican); Communication: Q1 (position 19 from 94, 1st Spanish and Iberoamerican); Journal Citation Indicator (JCI): Education: Q1 (position 12 from 739; 1st Spanish and Iberoamerican); Communication: Q1 (position 6 from 217, 1st Spanish and Iberoamerican).
SOCIAL SCIENCES CITATION INDEX: The top journal in Spanish in Communication since 2007.

SCOPUS

CITESCORE 2021 (2022-23): (9.8): Q1 in Cultural Studies (position 2nd from 1,127) (percentile 99). Q1 in Communication: position 9th from 467 (percentile 98). Q1 in Education (position 17th from 1,406) (percentile 98).
SCIMAGO JOURNAL RANK: SJR 2021 (2022-23): 1.382: Q1 in Cultural Studies, Communication and Education (first journal in Spanish language in Education, Communication and Cultural Studies).

RECYT (FECYT-MEC)

Ranking FECYT 2021 (2022-23): Education: 1st from 76 journals (99.82 points out of 100) (top 1%); Communication, Information and Scientific Documentation: 2nd from 24 journals (93.08 points out 100) (top 1%);
FECYT 2021 Seal of Quality Excellence (12 indicators).

GOOGLE SCHOLAR

2021 (2022-23): Top 100 of Google: Position 3rd (from 100) in the Spanish ranking of all research journals areas. H5: 44. Mediana H5: 58. En 2022-06-12: H: 102; H5: 82 (56,884 accumulated citations).
According EC3 Reports, 2018 (UGR): H5 Index (2013-2017); 1st in Education (out of 165; H5: 38); 1st in Communication (out of 51; H5: 38).

DIALNET METRICS

Q1: EDUCATION 2020 (2022): position 1st (from 230); IF: 4.545 (909 cites).
Q1: COMMUNICATION 2020 (2022): position 2nd (from 62); IF: 4.545 (909 cites).

CIRC (INTEGRATED CLASSIFICATION OF SCIENTIFIC JOURNALS) (EC3 Metrics)

2022, Level A+ (highest rated).

REDIB (CSIC) (IBERO-AMERICAN INNOVATION AND SCIENTIFIC KNOWLEDGE NETWORK)

2022: 2nd position among 1,199 selected journals from all areas. Rating: 58.364.

ERIH+

Level INT2 (2022).

PUBLISHED BY: GRUPO COMUNICAR EDICIONES
- www.comunicarjournal.com
- www.grupocomunicar.com

- Administration: info@grupocomunicar.com
- Staff: director@grupocomunicar.com
- Mail box 527. 21080 Huelva (Spain)

- COMUNICAR is a member of CEDRO (Spanish Centre for Reprographic Rights).

- COMUNICAR is a quarterly scientific publication, published in January, April, July and October.

- COMUNICAR journal accepts and promotes online institutional exchanges with other scientific journals.

INTERNATIONAL CO-EDITIONS

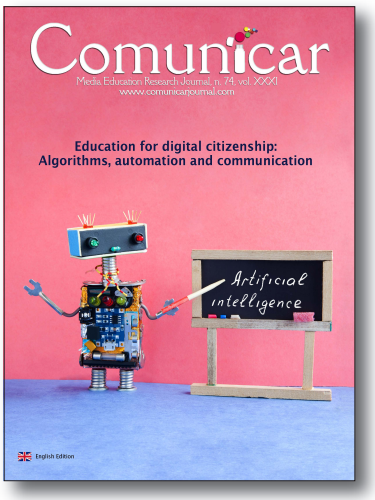
- **ENGLISH EDITION** (www.revistacomunicar.com/index.php?contenido=coeditores&idioma=en#EN)
- **PORTUGUESE EDITION** (www.revistacomunicar.com/index.php?contenido=coeditores&idioma=en#BR)
- **CHINESE EDITION** (www.revistacomunicar.com/index.php?contenido=coeditores&idioma=en#CN)
- **LATIN-AMERICAN EDITION** (www.revistacomunicar.com/index.php?contenido=coeditores&idioma=en#IB)
- **RUSSIAN EDITION** (www.revistacomunicar.com/index.php?contenido=coeditores&idioma=en#RU)

PRINTED BY: Estugraf. Madrid (Spain)

E-SHOP: www.revistacomunicar.com/index.php?contenido=tienda

CONTENTS

Comunicar, 74, XXXI (2023-1)



Education for digital citizenship: Algorithms, automation and communication

Educación para la ciudadanía digital: Algoritmos, automatización y comunicación

DOSSIER

THEMATIC EDITORS

Dr. Berta García-Orosa, University of Santiago de Compostela (Spain)
Dr. João Canavilhas, University of Beira Interior-Covilha (Portugal)
Dr. Suso Benítez-Baleato, Harvard University (USA)

KALEIDOSCOPE

© ISSN: 1134-3478 • e-ISSN: 1988-3293

SUBMISSION GUIDELINES

GENERAL INFORMATION

'Comunicar', Media Education Research Journal is published by Grupo Comunicar Ediciones (VAT: G21116603). This established non-profit professional group, founded in 1988 in Spain, specialises in the field of media education. The journal has been in print continuously since 1994, published every three months.

Contents are peer reviewed, in accordance with publication standards established in the APA 7 (American Psychological Association) manual. Compliance with these requirements facilitates indexation in the main databases of international journals in this field, which increases the dissemination of published papers and therefore raises the profile of the authors and their centres.

'Comunicar' is indexed in the Social Sciences Citation Index (SSCI), Journal Citation Reports (JCR), Scisearch, Scopus and over 790 databases, catalogues, search engines and international repertoires worldwide.

Each issue of the journal comes in a print (ISSN:134-3478) and electronic format (www.comunicarjournal.com) (e-ISSN: 1988-3293), identifying each submission with a DOI (Digital Object Identifier System).

SCOPE AND POLICY

Subject Matter: Fundamentally, research papers related to communication and education, and especially the intersection between the two fields: media education, educational media and resources, educational technology, IT and electronic resources, audiovisual, technologies... Reports, studies and experiments relating to these subjects are also accepted.

Contributions: 'Comunicar' publishes research results, studies, state-of-the-art articles and bibliographic reviews especially in relation to Latin America and Europe and regarding the convergence between education and communication, preferably written in Spanish although submissions are also accepted in English. The contributions to this journal may be: Research papers, Reports, Studies and Proposals (5,000-6,700 words of text, references included), State-of-the-art articles: (6,000-7,200 words of text, including references).

Unsolicited manuscripts sent in by authors are initially placed in the Miscellaneous section of the journal. The Topics section is organized by an editor through a system of call for papers and specific commissions to experts in the field. If we receive manuscripts within the deadline for a particular topic, the journal editor can refer the manuscript to the Topics editor for assessment and possible publication in this monographic section. The deadline for each Topic section is at least nine months before publication.

EDITORIAL PROCESS

'Comunicar' confirms receipt of all articles submitted by authors and keeps the authors informed of the acceptance/rejection process, as well as the editing process, in the event of acceptance. The webpage www.comunicarjournal.com also provides information about manuscript editing procedures, full guidelines for publication, as well as the Annexes: review protocol prior to submission, manuscript structure protocol, external review protocol…

In general, once the external reports have been reviewed, the criteria that justify the editors' decision to accept or reject submissions are as follows: a) Topical and new; b) Relevance: applicability of the results to the resolution of specific problems; c) Originality: valuable information, repetition of known results; d) Significance: advancement of scientific knowledge; e) Reliability and scientific validity: verified methodological quality; f) Organisation (logical coherence and material presentation); g) Presentation: good writting style.

PRESENTATION AND STRUCTURE OF ORIGINAL PAPERS

Manuscripts must be sent exclusively through the Journal Management Centre (www.revistacomunicar.com/ojs/). These publication guidelines are based on the standards of the American Psychological Association (APA 7): (https://bit.ly/3cN4XX9).

Structure: The following two files must be sent together: manuscript (main text) and cover letter.

ETHICAL COMMITMENT AND RESPONSIBILITIES

Each author must submit a statement of authorship and text originality. Previously published material will not be accepted. The cover letter must specify the transfer of copyright ownership of the manuscript for its publication in 'Comunicar'.

www.comunicarjournal.com

Comunicar©

MEDIA EDUCATION RESEARCH JOURNAL

XXXI, 74

© ISSN: 1134-3478 • e-ISSN: 1988-3293

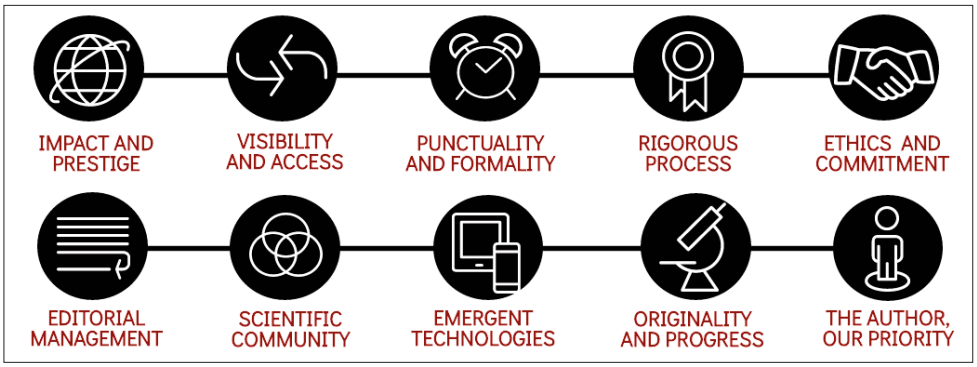
Comunicar©
Quality Criteria
Registered in the Spanish Patent and Trademark Office, with the code1806709

GLOBAL MIL AWARDS
UNESCO, 2019

MARIANO CEBRIÁN TOP PRIZE
University of Zaragoza/Aragon-Radio, 2015

TOP PRIZE IN COMMUNICATION
Carlos III University, Madrid, 2007

QUALITY CRITERIA



The quality criteria of 'Comunicar' Media Education Research Journal are the indicators of our commitment to a rigorous and professional publishing process. They constitute the endorsement for authors, reviewers, collaborators and readers that the final product offered meets the highest standards and academic rigour expected from a publication of national and international reputation.

• **Impact and prestige** are guaranteed by being in the first quartile among the most recognised journals: Q1 in Journal Citation Reports (JCR-SSCI) in Communication and Education; Q1 in Scopus (SJR) in Cultural Studies and Q1 in Scopus (CiteScore) in Communication and Education.

• **Visibility and access** are safeguarded by an open and shared access policy that makes all published manuscripts available to any reader in a bilingual version: Spanish and English.

• **Punctuality and formality** contribute to an efficient flow of manuscripts within established timeframes, facilitating quarterly publication thanks to highly effective schedule compliance.

• **A rigorous process** is supported by an International Reviewers Board of nearly 1,059 highly qualified researchers in the fields of education and communication from almost 54 countries all over the world.

• **Ethics and commitment** that ensure the upholding of rights and duties that protect the whole academic community: authors, reviewers, readers and editors. Our Code of Ethics is based on the International Committee on Publication Ethics (COPE), ensuring guidance for authors.

• **High quality Editorial Management** is accomplished throughout the OJS Platform, from the Science and Technology Foundation (FECYT), complemented by reviewers´ contributions and stylistic department work in English and Spanish languages.

• Our **scientific community** is fostered and increased by an influential academic community of authors (2,036), reviewers (1,059) and readers all over the world. This Community is supported by a committed editorial group composed of: an Editorial Board, an Academic Committee, an International Reviewers Board, an Advisory Board and a Technical Board.

• Use and development of **emergent technologies** encourages the diffusion and impact of the journal, continuously adjusting the access formats, communication models and academic platforms for academic diffusion (ResearchGate, AcademiaEdu...).

• The **originality and progress** of the outstanding research contributions from the research areas of communication and education are guaranteed by plagiarism controls (CrossCheck), applied to all manuscripts accepted before being published to preserve the originality and progress of research.

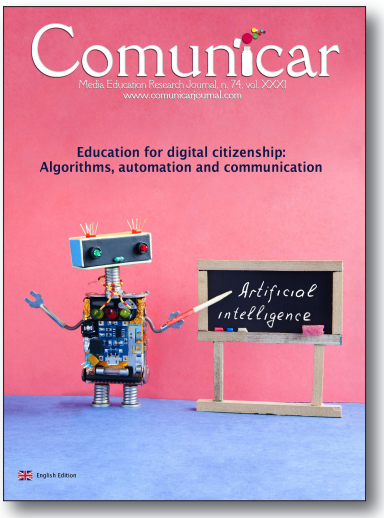
• At last but not least the **author is our priority,** since after all, authors uphold and give sense to entire process. Every manuscript is available in the Journal website with accurate information about citation, statistical data, references, impact metrics and interaction in social media.

Quality criteria are, in summary, a set of standards that guarantee the whole process, ensuring a professional treatment for every person involved in the publishing, reviewing, editing and spreading processes of the manuscripts.

Information on evaluators, acceptance/rejection rates and internationalisation in Comunicar 74

• Number of research works received: 216. Number of research works accepted: 10.
• Percent of manuscripts accepted: 4.63%; Percent of manuscript rejected: 95.37%.
• Received manuscripts internationalisation: 43 countries.
• Numbers of Reviews: 335 (109 internationals and 226 nationals) (update: www.comunicarjournal.com).
• Scientific Reviewers internationalisation: 29 countries.
• Country of origin: 7 countries (Germany, Egypt, Spain, Malaysia, Pakistan, Portugal & Romania).
• International databases in COMUNICAR 74: 811 (2023-1) (update: www.comunicarjournal.com).

© ISSN: 1134-3478 • e-ISSN: 1988-3293



Comunicar 74

Special issue

Education for digital citizenship: Algorithms, automation and communication

Comunicar

Clarivate Analytics
JOURNAL CITATION REPORTS

Journal Impact Factor (JIF) · 5.725
18th / 267 Education
19th / 94 Communication
Journal Citation Indicator (JCI)
12th / 739 Education
6th / 217 Communication

Scopus®

CiteScore · 9.8
2nd / 1,127 Cultural Studies
17th / 1,406 Communication
9th / 467 Education

© ISSN: 1134-3478 • e-ISSN: 1988-3293

Comunicar, n. 74, v. XXXI, 2023 | Media Education Research Journal | ISSN: 1134-3478; e-ISSN: 1988-3478
www.comunicarjournal.com

Algorithms and communication: A systematized literature review

Algoritmos y comunicación: Revisión sistematizada de la literatura

Dr. Berta García-Orosa. Professor, Department of Communication Sciences, Universidade de Santiago de Compostela (Spain) (berta.garcia@usc.es) (https://orcid.org/0000-0001-6126-7401)
Dr. João Canavilhas. Associate Professor, Department of Communication, Philosophy and Politics, Universidade da Beira Interior, Covilhã (Portugal) (jc@ubi.pt) (https://orcid.org/0000-0002-2394-5264)
Dr. Jorge Vázquez-Herrero. Assistant Professor, Department of Communication Sciences, Universidade de Santiago de Compostela (Spain) (jorge.vazquez@usc.es) (https://orcid.org/0000-0002-9081-3018)

ABSTRACT

The influence of algorithms on society is increasing due to their growing presence in all areas of daily life. Although we are not always aware of it, they sometimes usurp the identity of other social actors. The main purpose of this article is to address the meta-research on the field of artificial intelligence and communication from a holistic perspective that allows us to analyze the state of academic research, as well as the possible effects on these areas and on the democratic system. To this end, we carried out a systematized review of recent literature using quantitative and qualitative approaches. The subject analyzed is changing and novel: it includes the impact and interaction of algorithms, bots, automated processes, and artificial intelligence mechanisms in journalism and communication, as well as their effects on democracy. The results show expanding scientific production, mostly in English, based on theoretical discussion or focused on the perception of communication professionals. The object of study is centered mostly on journalism and democracy, and to a lesser degree on ethics or education. Studies indicate great interest in the effects of the use of algorithms on journalism and democracy, but the answers are still uncertain and the challenges for the coming years are significant.

RESUMEN

La influencia de los algoritmos en la sociedad es cada vez mayor a través de una presencia creciente en todos los ámbitos de la vida diaria, sin que seamos conscientes de ello y, en ocasiones, usurpando la identidad de otros actores sociales. El artículo tiene como propósito principal abordar la metainvestigación sobre el campo de la inteligencia artificial y la comunicación, desde una perspectiva holística que permita analizar el estado de la investigación académica, así como los posibles efectos en estas dos áreas y en la convivencia en un sistema democrático. Para ello se lleva a cabo una revisión sistematizada de la literatura reciente desde enfoques cuantitativos y cualitativos. La temática analizada es cambiante y novedosa; incluye el impacto y la interacción de algoritmos, bots, procesos automatizados y mecanismos de inteligencia artificial en el periodismo y la comunicación, así como su efecto en la democracia. Los resultados dibujan una producción científica en expansión, mayoritariamente en inglés, basada en la discusión teórica o centrada en la percepción de los profesionales de la comunicación. El objeto de estudio mayoritario se sitúa en el periodismo y en la democracia, con menor implicación de la ética o la educación. Los estudios señalan un gran interés sobre los efectos del uso de algoritmos sobre el periodismo y la democracia, pero las respuestas son todavía inciertas y los retos para los próximos años importantes.

KEYWORDS | PALABRAS CLAVE

Artificial intelligence, communication, journalism, democracy, public opinion, review.
Inteligencia artificial, comunicación, periodismo, democracia, opinión pública, revisión.

Received: 2022-04-28 | Reviewed: 2022-06-11 | Accepted: 2022-07-19 | OnlineFirst: 2022-10-30 | Published: 2023-01-01
DOI https://doi.org/10.3916/C74-2023-01 | Pages: 9-21

1. Introduction

Algorithms have become actors in the social, economic, political, and cultural spheres in recent years. Daily life and the decisions people make are increasingly tied to mathematical models and big data, "with varying degrees of opacity as to how they operate, in whose interest, and with what implications" (Thurman et al., 2019). Though at times algorithms may replace people's decision-making with software (Broussard et al., 2019), at other times they increase the commodification of audiences (García-Orosa, 2018), pre-designing so-called algorithmic audiences (Eldridge et al., 2019). "Algorithms have become a widespread trope for making sense of social life" (Ziewitz, 2017), and they have a greater capacity to shape the public sphere than at other times in their history (Broussard et al., 2019).

However, this situation does not exist in isolation; rather, it is part of a stage in digital communication characterized by events that are designed by the use of algorithms and that characterize the fourth wave of digital communication: digital platforms. These digital platforms have become actors in all phases of communication, the intensive use of artificial intelligence and big data, the uncritical use of technology, and the heightened striving for engagement with the audience, alongside three great challenges for democracy: a) polarization; b) fake news, deepfakes and astroturfing; and c) echo chambers and bubble filters (García-Orosa, 2022). This situation has led to noteworthy changes in the profession, in research, and in the teaching of journalism and communication, as well as in the public sphere and democratic society. The use of bots and artificial intelligence in political campaigns and referenda has been extensively studied in recent years (García-Orosa et al., 2021), with results that point not only to algorithms' direct influence on results but also towards a reconfiguration of the public sphere (Papakyriakopoulos et al., 2018; Helberger, 2019). Democracy will have to be reimagined in the new communication paradigm (Castells, 2022).

At the same time, the scientific community is taking on an object of study whose strength lies, in part, in the concealment of its functioning, identity, and objectives. The growing influence of algorithms in economic, political, social, and media systems in recent years has been accompanied by a skyrocketing increase in scientific research in those fields. We are witnessing a turning point, not only because of the changes that the pandemic has produced in communication and public opinion but, above all, because of the need to update research methods in order to make sense of an ever-changing object of study. Meta-analysis allows us to take a snapshot of scientific knowledge about an area and point out its shortcomings. In previous studies, such as the review of the scientific literature on communication in the Spanish-speaking world between 2013 and 2017 (Piñeiro-Naval & Morais, 2019), the issues addressed in this paper had yet to become influential. Currently, a literature review is needed to document milestones and forecast upcoming challenges. This article seeks to review scientific research on algorithms and communication from a holistic perspective that allows us to study their different uses in journalism and political and organizational communication, as well as their effects on these fields and democratic society. To that end, we conducted a quantitative and qualitative systematic review of recent literature.

2. Material and methods

To analyze the recent scientific research on the intersection of AI and the field of communication, specifically journalism, political and organizational communication, and democracy, we conducted a systematic literature review. The study employed a systematic, scientifically-rigorous approach in the gathering, evaluation, analysis, and synthesis of data (Grant & Booth, 2009). The main objective was to evaluate the current state of research on a changing and novel topic that includes the impact and interaction of algorithms, bots, automated processes, and artificial intelligence mechanisms in journalism, political communication and organizations, as well as their effect on democracy. This frame of reference will allow us to advance current knowledge and suggest future areas of research, based on the identification of trends, strengths, and weaknesses in published studies (Shahnazi & Afifi, 2017). We developed the following research questions:

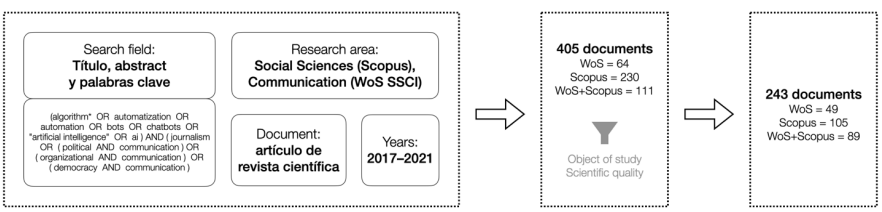
- RQ1. What are the characteristics of scientific research on artificial intelligence and communication?
- RQ2. What are the objects of study and methods employed within scientific research on artificial intelligence and communication?

- RQ3. What are the main areas of scientific research on artificial intelligence and communication?

Two scientific databases were used in the data collection phase: Web of Science (Clarivate Analytics) and Scopus (Elsevier). The selected articles include the terms in English, since the indexed publications' title, abstract, and keywords are in that language) and meet the conditions set forth in the search equation (Figure 1). The following additional inclusion criteria were considered: articles published in scientific journals, published between 2017 and 2021 (including some that were published online first), and in the categories of communication (Web of Science) and social sciences (Scopus).

The resulting set of documents consisted of 64 articles from Web of Science (SSCI), 230 articles from Scopus, and 111 articles found in both (in total, 405). In evaluating the dataset, the title, abstract, and methods were verified to apply a series of exclusion criteria based on adequacy and quality. First, we verified how each document deals with the object of study of this review, discarding the articles that did not deal with the relationship between artificial intelligence and the field of communication as defined in the search equation. Secondly, we made sure the articles met the standards of scientific rigor, though we also assumed they did because they are published in journals listed in the indicated databases. The final sample consists of 243 documents[1].

Figure 1. Synthesis of the systematized literature review process



The parameters for qualitative analysis of the selected documents are presented in Table 1, which represents the systematic categorization of each article as indicated by the review guidelines (Codina, 2018), implemented manually and by a single coder. Finally, a visualization of the Scopus results (n=194) for the analysis of co-citation and keywords in both databases is created with the VOSviewer software. However, due to limitations in the import and export of cited references, we were not able to combine the visualization of co-citation in WoS and Scopus. Subsequently, we performed a qualitative analysis of the documents and highlighted the areas analyzed within them.

Table 1. Article analysis variables in the systematic literature review

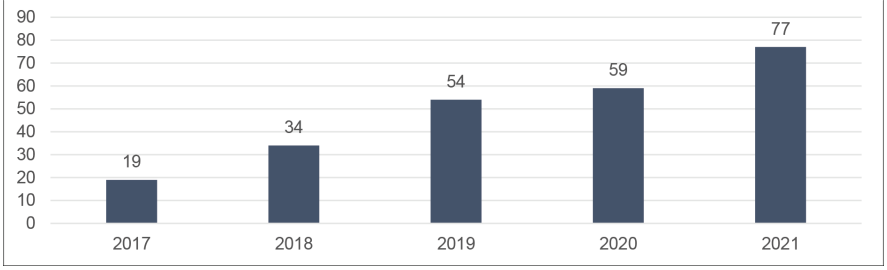
Article identification	Database
	DOI
	Authorship
	Number of authors
	First signatory
	Authors' gender
	Home country of affiliated institution
	Title
	Journal
	Volume, number, and pages
	Year of publication
Contents	Summary
	Keywords
	Language
	Open access
	Object of study
Methodology	Methods used
Impact	Citations
	Co-citation

3. Analysis and results

3.1. Bibliometric analysis

RQ1 (What are the characteristics of scientific research on artificial intelligence and communication?) was answered first. A substantial amount of research, based on 243 articles, was conducted on AI between 2017-2021. Over the years, interest has increased (Figure 2).

Figure 2. Scientific articles by year of publication



Articles on AI and communication were found in 97 different scientific journals, including, but not limited to, "Digital Journalism" (44 articles), "New Media & Society" (18), "Media and Communication" (12), "Journalism Practice" (10), and "Profesional de la Información" (9) (complete list in appendix[1]). Most of the articles are published exclusively in English (86.4%, Table 2). The presence of Spanish-language, Russian, and Slovenian journals in the impact indices ensures the publication of articles on this subject in other languages. Open access, through the journals themselves, is available for 58.4% of the articles.

Table 2. Articles published per language and number of authors

Language(s)	Articles	%
Slovenian	4	1.7
Spanish	12	4.9
Spanish and English	9	3.7
English	210	86.4
Portuguese	2	0.8
Portuguese and English	5	2.1
Russian	1	0.4
Authors	**Articles**	**%**
1	89	36.6
2	73	30.0
3	52	21.4
4	23	9.5
5	4	1.7
6	2	0.8

Nearly 37% percent of articles have only one author, while the majority have at least two (Table 2). Men make up 63.3% of authors, and the remaining authors are women. In 60.6% the lead author is male; however, this variable has error-inducing limitations in its coding, and in 1.4% of the articles it was not possible to determine the lead author's gender.

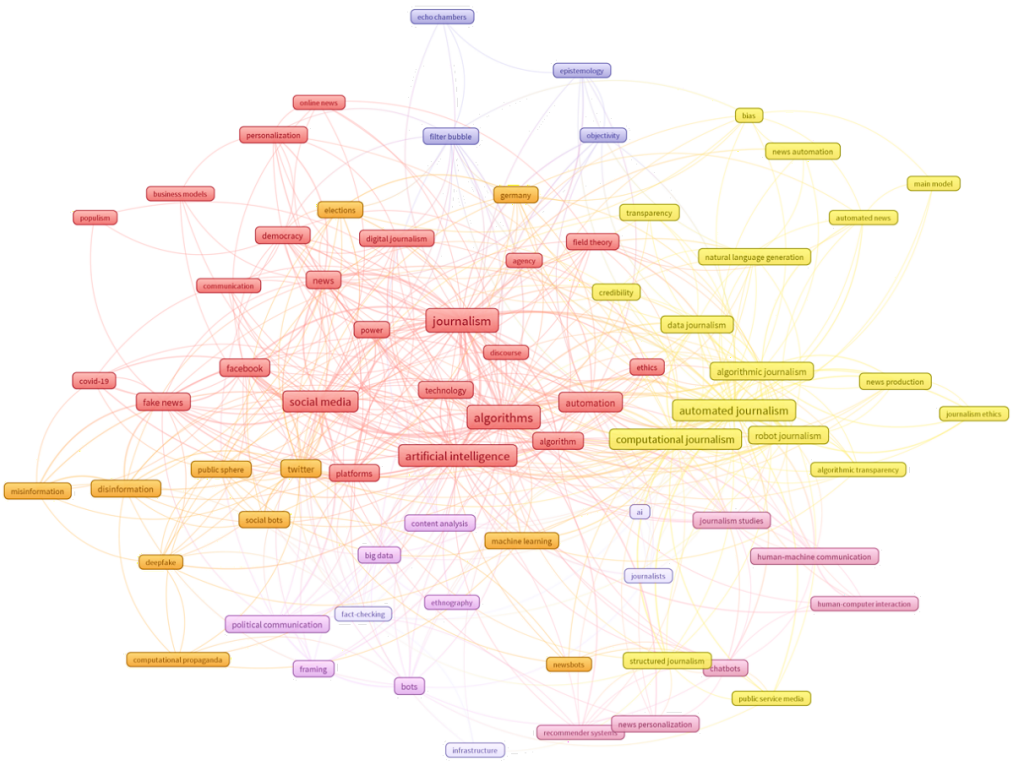
The main authors, among the 9,391 identified in the co-citation analysis carried out with the documents retrieved in Scopus (n=194), are Nicholas Diakopoulos (149), Seth C. Lewis (102), Matt Carlson (95), Neil Thurman (77), Philip N. Howard (76), Chris W. Anderson (74), Natali Helberg (72), Andreas Graefe (64), Rasmus Kleis Nielsen (63), and Nic Newman (61). Therefore, the preeminent authors in the study of AI's intersection with communication are largely from American and British universities. Appendix[2] presents the co-citation graph for authors with at least 20 citations, where the most frequently-cited authors make up a cluster colored in red.

The 515 authors of the articles analyzed are affiliated with institutions across 37 countries (appendix[3]). The top five are the United States (128; 24.9%), Spain (54; 10.5%), the United Kingdom (49; 9.5%), Germany (37; 7.2%), and the Netherlands (30; 5.8 %).

In the systematic literature review, we analyzed the articles' objects of study to answer RQ2 (What are the objects of study and methods of scientific research on artificial intelligence and communication?). A total of 844 keywords were assigned in the articles in both WoS and Science; from the keyword graph

(Figure 3), which represents the 73 terms with a frequency greater than or equal to three, we identified seven initial clusters. The largest group (depicted in red) covers journalism, AI, algorithms, platforms, and social media. A second cluster (depicted in yellow) represents automated, robotic, and computational journalism. The third cluster (in orange) corresponds to disinformation. Among the remaining clusters, the most significant represents the field of political communication.

Figure 3. Graph of keywords assigned by the authors to the articles



The keyword analysis is a first look at the object of study, which we complemented with a detailed analysis following a reading of the article. The results (Table 3) indicate that the most studied topic is the impact of artificial intelligence on journalism, taking into account its influence on news production, audiences, and the profession. Secondly, researchers studied AI's effects on the public sphere, democracy, and political communication.

To a lesser extent, researchers studied the connection between AI and web platforms, fundamentally those of social media. Other aspects analyzed are related to the rise of misinformation and fact-checking initiatives; scientific research itself, metascience and the agenda for future research; the impact on communication and organization management; ethical and regulatory issues; and education and digital literacy.

Regarding the methods used (Table 3), a significant number of articles (26.7%) focus on a theoretical-conceptual discussion without an explicit methodology, which can be added to the set of literature reviews (5.4% of articles) to form a group of theoretical articles. Researchers employed various methodologies in studying artificial intelligence in the field of communication, with a focus on the perception of practitioners, experts, and consumers in numerous studies, as seen in interviews, surveys, and focus groups (21.3%).

Data analysis methods (14.2%), generally applied to social networks, have a specific value due to the significant connection between platforms and algorithms. Content analysis, both quantitative and qualitative, is the fourth most commonly-employed methodology, followed by case studies. Other

methodological approaches appear less frequently, such as experiments, field work, methodological discussion, and participatory research.

Table 3. The articles' objects of study and methodologies

Object of study		Articles	%
Journalism		90	37.0
	Production	45	18.5
	Profession	16	6.6
	Audiences	29	11.9
Public sphere, democracy, and political communication		56	23.0
Platforms		44	18.1
Disinformation and fact-checking		15	6.2
Research		15	6.2
Communication and management in organizations		11	4.5
Ethics and regulation		8	3.3
Education and literacy		4	1.6
Method		**Articles**	**%**
Theoretical discussion (unspecified method)		64	26.7
Data analysis		34	14.2
Interviews		33	13.8
Content analysis		29	12.1
Case study		27	11.3
Experiment		21	8.8
Poll		16	6.7
Literature review		13	5.4
Field work		9	3.8
Others		9	3.8
Methodological discussion		5	2.1
Participatory research		3	1.3
Focus group		2	0.8

Note. The method could not be coded for three inaccessible articles.

The impact of the articles reviewed is limited by the short amount of time since publication, since the period covered is 2017-2021. Still, the articles are frequently cited, with an average of 12 citations each, in a total of 2,913. Appendix 4 lists the ten most frequently-cited articles according to the databases consulted, whose authors include 20 men and 7 women[4].

3.2. Qualitative interpretation: From news production to pre-constructed audiences

Algorithms have been used in scientific literature for data extraction and interpretation, especially in content analysis and experiments (Broersma & Harbers, 2018; Yarchi et al., 2020). However, in recent years, algorithms themselves have become an object of study, either because of their influence on some of the traditional actors in politics, economy, society, or culture, or because of their role as political actors.

This section is structured based on the qualitative results of the systematic review of the literature carried out to answer RQ3 (What are the main areas of scientific research on artificial intelligence and communication?). Research in recent years has centered on several main narratives: the influence of algorithms on democracy, the effects of algorithms on the media and audiences, and the significance of algorithmic determination of consumption. The following is a review of the approach and results of the studies conducted on these topics.

3.2.1. Influence on democracy

As noted in the introduction, the literature indicates that the widespread use of algorithms greatly influences the functioning of democratic political systems, and that bots' influence also continues to grow (Montal & Reich, 2017; Santini et al., 2018), especially during campaign season. Research on this topic focuses on a technical definition of algorithms and primarily seeks to develop detection systems through machine learning (Häring et al., 2018; Dubois & McKelvey, 2019). Meanwhile, in the social sciences, researchers question the health of democracy due to the spread of fake news (Bimber & Gil-de-Zúñiga, 2020), as well as astroturfing campaigns that can manipulate and sow uncertainty (Zerback et al., 2021).

Today, doubt has been cast on some concepts that were otherwise widely accepted in recent years, such as bubble filters (Puschmann, 2019). Some studies indicate that social media reinforce existing attitudes (Ohme, 2021). Others discuss social media's influence on the public sphere (Kaluža, 2021). There also are studies that question the validity of the term. Haramba et al. (2018) propose a historical interpretation from the perspective of the commodification of readership (García-Orosa, 2018). The goal of satisfying readers' habits, even with false, misleading, or biased information, stems from the attention

economy and is a fundamental principle of algorithms that satisfies users by captivating them (Seaver, 2018). In this sense, Schjøtt-Hansen and Hartley (2021) analyze algorithms and news selections to describe the evolution from distributing news content to readers/viewers treated as segments of consumer groups to algorithmically constructing individual readers/viewers as aggregate data points.

After a decade of euphoria about platforms' potential to empower citizens, disinformation, fake news, incitement to hatred and the Cambridge Analytica scandal, among others, have engendered mistrust (van-Dijck, 2020). Scientific literature highlights the risks in algorithms' potential to slander and the defenselessness of the media and citizens, as pointed out by Lewis et al. (2019). The authors draw attention to two relevant issues: the difficulty in finding guilty parties in defamation cases and in finding defenses as powerful as those wielded by the platforms.

These influences on democracy have led some authors to speak of "algorithmic culture" (Gilbert, 2018) and potential threats to democratic values. In recent years, scholars have called for legislative reforms to address the new challenges that online communication poses for democratic values or specific issues such as legislation on bots (Jones & Jones, 2019), privacy and facial recognition (Leong, 2019), or incidences of racism caused by algorithms (Turner-Lee, 2018).

3.2.2. Journalism and media

State-of-the-art technology has affected the practice of journalism in recent years (López-García & Vizoso, 2021), and the use of algorithms has sparked debates about the industry's core definition and foundation. Researchers have coined different names for the use of algorithms, (Vállez & Codina, 2018) among which automated, algorithmic, or robot journalism are the most used. Under this label, scholars have analyzed, fundamentally from the perspective of journalists and media directors themselves, the consequences of the implementation and use of algorithms in the production, distribution, and circulation of information.

In recent years, a growing number of media outlets, such as The Associated Press, The Washington Post, and the BBC, have embraced "automated journalism," (Graefe, 2018), also known as "algorithmic journalism" (Kotenidis & Veglis, 2021) or "robot journalism" (Waddell, 2018), understood as the automatic generation of journalistic texts through software and algorithms, with little or no human intervention, except for the initial programming (Danzon-Chambaud & Cornia, 2021; Sehl et al., 2021). Nonetheless, algorithms also intervene in the phase of selecting the issue, sources, and circulation of the journalistic message. Automation is studied from the perspective of helping journalists, for example, in the search for newsworthy events (Diakopoulos et al., 2021; Thurman et al., 2017), in personalized distribution by news recommendation systems (Helberger, 2019), in promoting data journalism (Tong & Zuo, 2021), in evaluating the credibility of sources (Fletcher et al., 2020; Graefe et al., 2018), or in redefining news values (Choi, 2019).

Overall, the results reveal the transformative role of machines, especially in the news-gathering and distribution phases, and increasingly in the writing phase, especially in data-rich specialties such as sports and economics. However, journalists continue to control all phases of the news production process (Milosavljevi & Vobi , 2019), especially in the news selection and editing phases, suggesting a desire to protect their role as final arbiters of meaning (Wu et al., 2019). Several authors have studied the potential for the robotization of journalism (Borges & Gambarato, 2019; Dierickx, 2021), and some have concluded that robots do not threaten their work (De-la-Torre, 2020).

As changes in the profession come to light, a significant re-working of the logic of journalism is leading to a new conceptualization of the field and technology's influence on it. The studies examine automation as one element of journalists' work (Calvo-Rubio & Ufarte-Ruiz, 2020) and identify contradictions between automation and some of the fundamental ideals of journalism, like public service, autonomy and objectivity (Milosavljevi & Vobi , 2019), which leads to friction when implemented in newsrooms (Hermida & Young, 2017). Journalists point to the nature of the sources and robots' lack of a "nose for news" as some of the limitations of automated journalism (Thurman et al., 2017).

After the period of 2015 to 2016, which was partly characterized by a very favorable and uncritical attitude, the most recently published texts (2017-2019) once again opted for the neutral tone typical

of 2011 and 2012 (Parratt-Fernández et al., 2021). Research has emerged that questions the role of journalism within society and the consequences of algorithmizing the profession, as well as social, economic, political, and cultural life, and attempts to re-imagine the field (Bucher, 2017) and analyze challenges centered on ethics (Dörr & Hollnbuchner, 2017) and credibility (Tandoc et al., 2020). Such studies describe algorithmic journalism and new challenges in the fight against the dominance of commercial interests (more visible in outlets' business departments than in newsrooms) in the implementation of automation (Sla ek-Brlek et al., 2017).

Other authors highlight the growing dependence on software providers and platforms in the face of editorial independence, which should prevail in journalism (Schapals & Porlezza, 2020; Weber & Kosterich, 2018), due to the role of professionals, such as technologists, i.e., computer scientists or "technoactors," new to the (Canavilhas et al., 2016) production processes, who influence the news and redefine journalism with their practices and values (Wu et al., 2019).

3.2.3. Audience

Automation has sparked new debates on the production of journalistic texts and their authorship (Montal & Reich, 2017), and in some situations it is no longer possible to determine who produces the news (Wölker & Powell, 2021). Moreover, automation has also changed journalists' relationships with the audience, for example, through the use of newsbots as mediators between journalist and audience (Ford & Hutchinson, 2019). Since the beginning of online journalism, the audience has been part of the journalist's work (García-Orosa, 2018), but the use of algorithms is a step forward that has two implications. First, through audience monitoring, "[...] journalists can—and do—monitor social network users and their content via sophisticated, professional apps that are also used by police and security forces. (Thurman, 2018: 1). Secondly, journalists can create algorithmic audiences in line with the interests of news outlets. Martin (2021) warns of the risks of the mediatization of news visibility through opaque algorithms, as well as through the platformization of news (van-Dijck et al., 2018) and the metrification of news values.

Algorithms not only influence what content is featured; the audience is also ranked according to their interest in the platform. Regarding Facebook, Thorson et al. (2021) suggest that people who are algorithmically categorized as interested in news or politics are more likely to attract content to their feeds, regardless of their self-reported interest in civic content. In this sense, Papakyriakopoulos et al. (2018) discuss the relevance of hyperactive users (users with above-average activity on the network) in shaping public opinion and democracy. The authors study their influence, which affects public opinion on social networks, and warn of the possible adverse consequences of algorithms and recommendation systems for political systems. Therefore, one of the most important aspects is circulation. Media outlets have gone from disseminating content to audiences and managing their activities, to transforming the audience into constructors of the discourse and creating algorithmic audiences based on previously-obtained big data. Bodó (2019) describes how European media, instead of focusing on increasing user engagement in the short term, try to personalize the news to increase audience loyalty in the long term. "Unlike the 'platform logic of personalization', which uses personalization to produce engagement and sell audiences to advertisers, they have developed a "news logic of personalization" that uses personalization to sell news to audiences. (Bodó, 2019: 1054).

New social, political, and media roles are conquering spaces as algorithms, a generalized trope to give meaning to social life (Ziewitz, 2017), which not only shapes the agenda, but also constructs the audience (Thorson et al., 2021). The media seek an audience that is "constructed" rather than a naturally arising one (Eldridge et al., 2019). As such, algorithmic audiences are programmed (Møller-Hartley et al., 2021) to promote a "particularly populist 'profitable and normal' media experience" (Harper, 2017). Users are often defenseless because they are unaware of how news are filtered and prioritized (Powers, 2017) and how the user profile is predicted.

3.2.4. Algorithmic determination of consumption

Literature highlights recommendation systems as shapers of public opinion and, therefore, of civic participation in public life. The massive consumption of information on social media platforms, which has

dealt a blow to traditional media, has led to a significant dependence on the algorithmic determination of news consumption based on previous audience behavior, analyzed through big data, and possible distortions such as polarization (Shmargad & Klar, 2020). News personalization systems are viewed as black boxes that indicate a significant disconnection between the practice and theory of algorithmic transparency, particularly in non-community contexts (Bastian et al., 2021). The use of different data sources to predict what content will be interesting to readers raises concerns about possible audience fragmentation (Makhortykh & Wijermars, 2021); after tracking news personalization for six years and detecting platform commodification, Kunert and Thurman (2019) also raised concerns about data protection and the effects of recommendation systems.

But there are also traditional media projects that use news recommendation systems transparently to combat disinformation and create a European public sphere, which seems to be confirmed by an analysis of the news they have produced (Canavilhas, 2022). Such is the case of the European Broadcasting Union's "A European Perspective," whose PEACH ecosystem seeks to offer the most appropriate content to each user at the most opportune time and on the most appropriate device. The system highlighted by recent academic literature sparks an important debate on the opacity of recommendation and content adaptation systems and, therefore, on their role in democratic systems (Helberger, 2019).

4. Discussion and conclusions

In a fluid and hybrid context, algorithms stand out as new actors in communication and political, economic, and social systems. Their influence, often based on the use of confidential personal data or the concealment or theft of digital identities, has increased in recent years, resulting in more and more disinformation campaigns that use algorithms and bots to achieve a greater and faster impact.

News organizations have adapted in various ways to a digital media environment dominated by algorithmic gatekeepers like search engines and social media (Graves & Anderson, 2020). Communicative robots are defined as autonomously operating systems designed for the purpose of quasi-communicating with humans to enable other algorithm-based functionalities, often based on artificial intelligence such as Siri or Alexa (Hepp, 2020).

Quantitatively, scientific research on the intersection of artificial intelligence and communication increased significantly from 2017-2021. Most articles are published in English and have several authors. The United States, Spain, and the United Kingdom have the greatest presence in our review. The objects of study address the different perspectives of these two interacting fields, though the most common issues are the field of journalism, whether in terms of production, the profession itself, or the audiences; the impact on the public sphere, democracy, and political communication; and the role of algorithms on platforms. Methodologically speaking, researchers have employed a range of methods and techniques to study the phenomenon at hand, including but not limited to, theoretical-conceptual discussions without an explicit methodology; studying the perspective of key players; and analyzing data obtained from platforms.

From a qualitative point of view, the scientific literature on algorithms and communication describes an uncertain situation that is difficult to analyze due to algorithms' typical lack of transparency. Researchers addressed how algorithms work from an engineering and computer science standpoint, and showed their concern about how journalism implements algorithms as well as the effects on audiences and democracy. The results must be confirmed with future research on how different figures in democracy are enhanced or assisted, taking culture into account, among other factors (Jamil, 2021).

There will be myriad challenges in the coming years. Below are some that our analysis has revealed:

- The search for specific methodologies and analytical methods that allow us to understand a changing and opaque reality.
- Promotion of multidisciplinary research.
- Empirical studies on the effects of using algorithms in different systems.
- Promotion of comparative analyses between different countries that advance the state of knowledge through generalizable data.

5. Limitations

This is a literature review of research that already has its own epistemic and methodological biases. The search formula leads to limited results; we had to limit the field to the intersection of artificial intelligence with journalism, political communication, organizational communication, and democracy, because the inclusion of the term "communication" interfered with the data. In addition, because artificial intelligence is inherently opaque, the narrative espoused by key players in the media becomes salient, with the validity and bias that this implies.

Notes

[1]Dataset available at: https://doi.org/10.6084/m9.figshare.19411187.
[2]Graph of co-citation of bibliographic references per author: https://doi.org/10.6084/m9.figshare.19632741.v1.
[3]Map of authors by country of affiliated institution: https://doi.org/10.6084/m9.figshare.19632759.
[4]List of the ten most cited articles of the systematized literature review: https://doi.org/10.6084/m9.figshare.19632762.

Authors' Contribution

Idea, B.G.O., J.C; Literature review (state of the art), B.G.O; Methodology, J.V.H., B.G.O; Data analysis, J.V.H; Results, J.V.H., B.G.O; Discussion and conclusions, B.G.O., J.C., J.V.H; First draft, B.G.O., J.V.H; Final revisions, B.G.O., J.C., J.V.H; Project design and sponsorships, B.G.O.

Funding Agency

This research has supported by Radón en España: percepción de la opinión pública, agenda mediática y comunicación del riesgo – RAPAC (SUBV-13/2021; Consejo de Seguridad Nuclear).

References

Bastian, M., Helberger, N., & Makhortykh, M. (2021). Safeguarding the journalistic DNA: Attitudes towards the role of professional values in algorithmic news recommender designs. *Digital Journalism*, *9*(6), 835-863. https://doi.org/10.1080/21670811.2021.1912622

Bimber, B., & Gil-De-Zúñiga, H. (2020). The unedited public sphere. *New Media and Society*, *22*(4), 700-715. https://doi.org/10.1177/1461444819893980

Bodó, B. (2019). Selling news to audiences - A qualitative inquiry into the emerging logics of algorithmic news personalization in European quality news media. *Digital Journalism*, *7*(8), 1054-1075. https://doi.org/10.1080/21670811.2019.1624185

Borges, P.M., & Gambarato, R.R. (2019). The role of beliefs and behavior on Facebook: A semiotic approach to algorithms, fake news, and transmedia journalism. *International Journal of Communication*, *13*, 603-618. https://bit.ly/3MgzAp6

Broersma, M., & Harbers, F. (2018). Exploring machine learning to study the long-term transformation of news: Digital newspaper archives, journalism history, and algorithmic transparency. *Digital Journalism*, *6*(9), 1150-1164. https://doi.org/10.1080/21670811.2018.1513337

Broussard, M., Diakopoulos, N., Guzman, A.L., Abebe, R., Dupagne, M., & Chuan, C.H. (2019). Artificial intelligence and journalism. *Journalism and Mass Communication Quarterly*, *96*(3), 673-695. https://doi.org/10.1177/1077699019859901

Bucher, T. (2017). 'Machines don't have instincts': Articulating the computational in journalism. *New Media and Society*, *19*(6), 918-933. https://doi.org/10.1177/1461444815624182

Calvo-Rubio, L.M., & Ufarte-Ruiz, M.J. (2020). Perception of teachers, students, innovation managers and journalists about the use of artificial intelligence in journalism. *Profesional de la Información*, (1), 29-29. https://doi.org/10.3145/epi.2020.ene.09

Canavilhas, J. (2022). Artificial intelligence in journalism: Automatic translation and recommendation system in the project "A European Perspective" (EBU). *Revista Latina de Comunicación Social*, *80*, 1-13.

Canavilhas, J., Satuf, I., Luna, D., Torres, V., Baccin, A., & Marques, A. (2016). Jornalistas e tecnoatores: A negociação de culturas profissionais em redações on-line. *Revista Famecos*, (3), 23-23.

Castells, M. (2022). Digital politics: A paradigm shift. In B. García-Orosa (Ed.), *Digital political communication strategies. Multidisciplinary reflections* (pp. 5-7). Palgrave. https://doi.org/10.1007/978-3-030-81568-4

Choi, S. (2019). An exploratory approach to the computational quantification of journalistic values. *Online Information Review*, *43*(1), 133-148. https://doi.org/10.1108/OIR-03-2018-0090

Codina, L. (2018). *Revisiones bibliográficas sistematizadas: Procedimientos generales y framework para ciencias humanas y sociales*. [Master dissertation, Universitat Pompeu Fabra]. https://bit.ly/3L7vMGz

Danzon-Chambaud, S., & Cornia, A. (2021). Changing or reinforcing the "rules of the game": A field theory perspective on the impacts of automated journalism on media practitioners. *Journalism Practice*. https://doi.org/10.1080/17512786.2021.1919179

De-La-Torre, J. (2020). Los periodistas no creen que la Inteligencia Artificial pueda substituirlos. *Escudo digital*. https://bit.ly/3L7mDht

Diakopoulos, N., Trielli, D., & Lee, G. (2021). Towards understanding and supporting journalistic practices using semi-automated news discovery tools. In J. Nichols (Ed.), *Proceedings of the ACM on Human-Computer Interaction* (pp. 1-30). Association for Computing Machinery. https://doi.org/10.1145/3479550

Dierickx, L. (2021). News automation, materialities, and the remix of an editorial process. *Journalism*. https://doi.org/10.1177/14648849211023872

Dörr, K.N., & Hollnbuchner, K. (2017). Ethical challenges of algorithmic journalism. *Digital Journalism*, *5*(4), 404-419. https://doi.org/10.1080/21670811.2016.1167612

Dubois, E., & Mckelvey, F. (2019). Political bots: Disrupting Canada's democracy. *Canadian Journal of Communication*, *44*(2), 27-34. https://doi.org/10.22230/cjc.2019v44n2a3511

Eldridge, S.A., García-Carretero, L., & Broersma, M. (2019). Disintermediation in social networks: Conceptualizing political actors' construction of publics on Twitter. *Media and Communication*, *7*, 271-285. https://doi.org/10.17645/mac.v7i1.1825

Fletcher, R., Schifferes, S., & Thurman, N. (2020). Building the 'Truthmeter': Training algorithms to help journalists assess the credibility of social media sources. *Convergence*, *26*(1), 19-34. https://doi.org/10.1177/1354856517714955

Ford, H., & Hutchinson, J. (2019). Newsbots that mediate journalist and audience relationships. *Digital Journalism*, *7*(8), 1013-1031. https://doi.org/10.1080/21670811.2019.1626752

García-Orosa, B. (2018). Perfil de la audiencia de cibermedios: representación discursiva y praxis del receptor 2.0. *Palabra Clave*, *21*, 111-133. https://doi.org/10.5294/pacla.2018.21.1.6

García-Orosa, B. (2022). Digital political communication: Hybrid intelligence, algorithms, automation and disinformation in the fourth wave. In B. García-Orosa (Ed.), *Digital Political Communication Strategies*. Palgrave. https://doi.org/10.1007/978-3-030-81568-4

García-Orosa, B., Gamallo, P., Martín-Rodilla, P., & Martínez-Castaño, R. (2021). Hybrid intelligence strategies for identifying, classifying and analyzing political bots. *Social Sciences*, *10*(10). https://doi.org/10.3390/socsci10100357

Gilbert, A.S. (2018). Algorithmic culture and the colonization of life-worlds. *Thesis Eleven*, *146*, 87-96. https://doi.org/10.1177/0725513618776699

Graefe, A., Haim, M., Haarmann, B., & Brosius, H.B. (2018). Readers' perception of computer-generated news: Credibility, expertise, and readability. *Journalism*, *19*(5), 595-610. https://doi.org/10.1177/1464884916641269

Grant, M.J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, *26*, 91-108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Graves, L., & Anderson, C.W. (2020). Discipline and promote: Building infrastructure and managing algorithms in a "structured journalism" project by professional fact-checking groups. *New Media and Society*, *22*(2), 342-360. https://doi.org/10.1177/1461444819856916

Häring, M., Loosen, W., & Maalej, W. (2018). Who is addressed in this comment? Automatically classifying meta-comments in news comments. In K. Karahalios, A. Monroy-Hernández, A. Lampinen, & G. Fitzpatrick (Eds.), *Proceedings of the ACM on Human-Computer Interaction* (pp. 1-20). https://doi.org/10.1145/3274336

Harper, T. (2017). The big data public and its problems: Big data and the structural transformation of the public sphere. *New Media and Society*, *19*(9), 1424-1439. https://doi.org/10.1177/1461444816642167

Helberger, N. (2019). On the democratic role of news recommenders. *Digital Journalism*, *7*(8), 993-1012. https://doi.org/10.1080/21670811.2019.1623700

Hepp, A. (2020). Artificial companions, social bots and work bots: communicative robots as research objects of media and communication studies. *Media, Culture and Society*, *42*(7-8), 1410-1426. https://doi.org/10.1177/0163443720916412

Hermida, A., & Young, M.L. (2017). Finding the data unicorn: A hierarchy of hybridity in data and computational journalism. *Digital Journalism*, *5*(2), 159-176. https://doi.org/10.1080/21670811.2016.1162663

Jamil, S. (2021). Automated journalism and the freedom of media: Understanding legal and ethical implications in competitive authoritarian regime. *Journalism Practice*. https://doi.org/10.1080/17512786.2021.1981148

Jones, R., & Jones, B. (2019). Atomising the news: The (in)flexibility of structured journalism. *Digital Journalism*, *7*(8), 1157-1179. https://doi.org/10.1080/21670811.2019.1609372

Kaluža, J. (2021). Habitual generation of filter bubbles: Why is algorithmic personalisation problematic for the democratic public sphere? *Javnost*. https://doi.org/10.1080/13183222.2021.2003052

Kotenidis, E., & Veglis, A. (2021). Algorithmic journalism-Current Applications and Future Perspectives. *Journalism and Media*, *2*(2), 244-257. https://doi.org/10.3390/journalmedia2020014

Kunert, J., & Thurman, N. (2019). The form of content personalisation at mainstream, transatlantic news outlets. *Journalism Practice*, *13*(7), 759-780. https://doi.org/10.1080/17512786.2019.1567271

Leong, B. (2019). Facial recognition and the future of privacy: I always feel like… somebody's watching me. *Bulletin of the Atomic Scientists*, *75*(3), 109-115. https://doi.org/10.1080/00963402.2019.1604886

Lewis, S.C., Guzman, A.L., & Schmidt, T.R. (2019). Automation, journalism, and human-machine communication: Rethinking roles and relationships of humans and machines in news. *Digital Journalism*, *7*(4), 409-427. https://doi.org/10.1080/21670811.2019.1577147

López-García, X., & Vizoso, A. (2021). Periodismo de alta tecnología: Signo de los tiempos digitales del tercer milenio. *Profesional de la Información*, (pp. 30-30). https://doi.org/10.3145/epi.2021.may.01

Makhortykh, M., & Wijermars, M. (2021). Can filter bubbles protect information freedom? Discussions of algorithmic news recommenders in Eastern Europe. *Digital Journalism*. https://doi.org/10.1080/21670811.2021.1970601

Martin, F.R. (2021). Visibility, connectivity, agency: Journalism's prospects in an age of automated social news sharing. *Digigal Journalism*, *9*(8), 1190-1198. https://doi.org/10.1080/21670811.2021.1977670

Milosavljevi , M., & Vobi , I. (2019). Human still in the loop: Editors reconsider the ideals of professional journalism through automation. *Digital Journalism*, *7*(8), 1098-1116. https://doi.org/10.1080/21670811.2019.1601576

Møller-Hartley, J., Bengtsson, M., Hansen, A.S., & Sivertsen, M.F. (2021). Researching publics in datafied societies: Insights from four approaches to the concept of 'publics' and a (hybrid) research agenda. *New Media and Society*.

https://doi.org/10.1177/14614448211021045
Montal, T., & Reich, Z. (2017). I, robot. You, journalist. Who is the author? Authorship, bylines and full disclosure in automated journalism. *Digital Journalism*, *5*(7), 829-849. https://doi.org/10.1080/21670811.2016.1209083
Ohme, J. (2021). Algorithmic social media use and its relationship to attitude reinforcement and issue-specific political participation-The case of the 2015 European immigration movements. *Journal of Information Technology and Politics*, *18*(1), 36-54. https://doi.org/10.1080/19331681.2020.1805085
Papakyriakopoulos, O., Hegelich, S., Shahrezaye, M., & Serrano, J.C.M. (2018). Social media and microtargeting: Political data processing and the consequences for Germany. *Big Data and Society*, *5*(2), 1-15. https://doi.org/10.1177/2053951718811844
Parratt-Fernández, S., Mayoral-Sánchez, J., & Mera-Fernández, M. (2021). The application of artificial intelligence to journalism: An analysis of academic production. *Profesional de la Información*, (pp. 30-30). https://doi.org/10.3145/epi.2021.may.17
Piñeiro-Naval, V., & Morais, R. (2019). Study of the academic production on communication in Spain and Latin America. [Estudio de la producción académica sobre comunicación en España e Hispanoamérica]. *Comunicar*, *61*, 113-123. https://doi.org/10.3916/C61-2019-10
Powers, E. (2017). My news feed is filtered? Awareness of news personalization among college students. *Digital Journalism*, *5*(10), 1315-1335. https://doi.org/10.1080/21670811.2017.1286943
Puschmann, C. (2019). Beyond the bubble: Assessing the diversity of political search results. *Digital Journalism*, *7*(6), 824-843. https://doi.org/10.1080/21670811.2018.1539626
Santini, R.M., Agostini, L., Barros, C.E., Carvalho, D., De-Rezende, R.C., Salles, D.G., Seto, K., Terra, C., & Tucci, G. (2018). Software power as soft power: A literature review on computational propaganda effects in public opinion and political process. *Partecipazione e Conflitto*, *11*, 332-360. https://doi.org/10.1285/i20356609v11i2p332
Schapals, A.K., & Porlezza, C. (2020). Assistance or resistance? Evaluating the intersection of automated journalism and journalistic role conceptions. *Media and Communication*, *8*(3), 16-26. https://doi.org/10.17645/mac.v8i3.3054
Schjøtt-Hansen, A., & Hartley, J.M. (2021). Designing what's news: An ethnography of a personalization algorithm and the data-driven (re)assembling of the news. *Digital Journalism*. https://doi.org/10.1080/21670811.2021.1988861
Seaver, N. (2018). Captivating algorithms: Recommender systems as traps. *Journal of Material Culture*, *24*(4), 421-436. https://doi.org/10.1177/1359183518820366
Sehl, A., Cornia, A., & Nielsen, R.K. (2021). How do funding models and organizational legacy shape news organizations' social media strategies? A comparison of public service and private sector news media in six countries. *Digital Journalism*. https://doi.org/10.1080/21670811.2021.1968920
Shahnazi, A.F., & Afifi, T.F. (2017). Strategies for literature reviews. In M. Allen (Ed.), *The SAGE Encyclopedia of Communication Research Methods*. https://doi.org/10.4135/9781483381411
Shmargad, Y., & Klar, S. (2020). Sorting the news: how ranking by popularity polarizes our politics. *Political Communication*, *37*(3), 423-446. https://doi.org/10.1080/10584609.2020.1713267
Sla ek-Brlek, S., Smrke, J., & Vobi , I. (2017). Engineering technologies for journalism in the digital age: A case study. *Digital Journalism*, *5*(8), 1025-1043. https://doi.org/10.1080/21670811.2017.1338526
Tandoc, E.C., Yao, L.J., & Wu, S. (2020). Man vs. machine? The impact of algorithm authorship on news credibility. *Digital Journalism*, *8*(4), 548-562. https://doi.org/10.1080/21670811.2020.1762102
Thorson, K., Cotter, K., Medeiros, M., & Pak, C. (2021). Algorithmic inference, political interest, and exposure to news and politics on Facebook. *Information Communication and Society*, *24*(2), 183-200. https://doi.org/10.1080/1369118X.2019.1642934
Thurman, N. (2018). Social media, surveillance, and news work: On the apps promising journalists a 'crystal ball. *Digital Journalism*, *6*(1), 76-97. https://doi.org/10.1080/21670811.2017.1345318
Thurman, N., Dörr, K., & Kunert, J. (2017). When reporters get hands-on with robo-writing: Professionals consider automated journalism's capabilities and consequences. *Digital Journalism*, *5*(10), 1240-1259. https://doi.org/10.1080/21670811.2017.1289819
Thurman, N., Lewis, S.C., & Kunert, J. (2019). Algorithms, automation, and news. *Digital Journalism*, *7*(8), 980-992. https://doi.org/10.1080/21670811.2019.1685395
Tong, J., & Zuo, L. (2021). The inapplicability of objectivity: Understanding the work of data journalism. *Journalism Practice*, *15*(2), 153-169. https://doi.org/10.1080/17512786.2019.1698974
Turner-Lee, N. (2018). Detecting racial bias in algorithms and machine learning. *Journal of Information, Communication and Ethics in Society*, *16*(3), 252-260. https://doi.org/10.1108/JICES-06-2018-0056
Vállez, M., & Codina, L. (2018). Periodismo computacional: evolución, casos y herramientas. *Profesional de la Información*, *27*(4), 759-768. https://doi.org/10.3145/epi.2018.jul.05
Van-Dijck, J. (2020). Governing digital societies: Private platforms, public values. *Computer Law and Security Review*, *36*, 105377-105377. https://doi.org/10.1016/j.clsr.2019.105377
Van-Dijck, J., Poell, T., & De-Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford University Press. https://doi.org/10.1093/oso/9780190889760.001.0001
Waddell, T.F. (2018). A robot wrote this? How perceived machine authorship affects news credibility. *Digital Journalism*, *6*(2), 236-255. https://doi.org/10.1080/21670811.2017.1384319
Weber, M.S., & Kosterich, A. (2018). Coding the news: The role of computer code in filtering and distributing news. *Digital Journalism*, *6*(3), 310-329. https://doi.org/10.1080/21670811.2017.1366865
Wölker, A., & Powell, T.E. (2021). Algorithms in the newsroom? News readers' perceived credibility and selection of automated journalism. *Journalism*, *22*(1), 86-103. https://doi.org/10.1177/1464884918757072

Wu, S., Tandoc, E.C., & Salmon, C.T. (2019). Journalism reconfigured: Assessing human-machine relations and the autonomous power of automation in news production. *Journalism Studies*, *20*(10), 1440-1457. https://doi.org/10.1080/1461670X.2018.1521299

Yarchi, M., Baden, C., & Kligler-Vilenchik, N. (2020). Political polarization on the digital sphere: A cross-platform, over-time analysis of interactional, positional, and affective polarization on social media. *Political Communication*, *38*, 98-139. https://doi.org/10.1080/10584609.2020.1785067

Zerback, T., Töpfl, F., & Knöpfle, M. (2021). The disconcerting potential of online disinformation: Persuasive effects of astroturfing comments and three strategies for inoculation against them. *New Media and Society*, *23*(5), 1080-1098. https://doi.org/10.1177/1461444820908530

Ziewitz, M. (2017). A not quite random walk: Experimenting with the ethnomethods of the algorithm. *Big Data and Society*, *4*(2). https://doi.org/10.1177/2053951717738105

© ISSN: 1134-3478 • e-ISSN: 1988-3293

Comunicar, n. 74, v. XXXI, 2023 | Media Education Research Journal | ISSN: 1134-3478; e-ISSN: 1988-3478
www.comunicarjournal.com

How to automate the extraction and analysis of information for educational purposes

Cómo automatizar la extracción y análisis de información sobre ciencia ciudadana con propósitos educativos

Miriam Calvera-Isabal. Research Assistant, Research Group of Interactive and Distributed Technologies for Education (TIDE), University of Pompeu Fabra, Barcelona (Spain) (miriam.calvera@upf.edu) (https://orcid.org/0000-0003-4117-6953)
Dr. Patricia Santos. Researcher, Research Group of Interactive and Distributed Technologies for Education (TIDE), University of Pompeu Fabra, Barcelona (Spain) (patricia.santos@upf.edu) (https://orcid.org/0000-0002-7337-2388)
Dr. H.-Ulrich Hoppe. Full Professor, Rhine-Ruhr Institute for Applied Systems Innovation, Duisburg (Germany) (uh@rias-institute.de) (https://orcid.org/0000-0003-3240-5785)
Cleo Schulten. Research Assistant, Rhine-Ruhr Institute for Applied Systems Innovation, Duisburg (Germany) (cs@rias-institute.de) (https://orcid.org/0000-0003-3082-6084)

ABSTRACT

There is an increasing interest and growing practice in Citizen Science (CS) that goes along with the usage of websites for communication as well as for capturing and processing data and materials. From an educational perspective, it is expected that by integrating information about CS in a formal educational setting, it will inspire teachers to create learning activities. This is an interesting case for using bots to automate the process of data extraction from online CS platforms to better understand its use in educational contexts. Although this information is publicly available, it has to follow GDPR rules. This paper aims to explain (1) how CS communicates and is promoted on websites, (2) how web scraping methods and anonymization techniques have been designed, developed and applied to collect information from online sources and (3) how these data could be used for educational purposes. After the analysis of 72 websites, some of the results obtained show that only 24.8% includes detailed information about the CS project and 48.61% includes information about educational purposes or materials.

RESUMEN

El interés y la práctica de la ciencia ciudadana (CC) ha aumentado en los últimos años. Esto ha derivado en el uso de páginas web como herramienta de comunicación, recolección o análisis datos o repositorio materiales y recursos. Desde una perspectiva educativa, se espera que al integrar información sobre proyectos de CC en un entorno educativo formal, se inspire a los maestros a crear actividades de aprendizaje. Este, es un caso interesante para usar bots que automaticen el proceso de extracción de datos de webs de CC que ayuden a comprender mejor su uso en contextos educativos. Aunque esta información está disponible públicamente, se deben seguir las reglas de la ley de protección de datos o GDPR. Este artículo tiene como objetivo explicar: 1) cómo la CC se comunica y promueve en los sitios web; 2) cómo se diseñan, desarrollan y aplican los métodos de web scraping y las técnicas de anonimización para recopilar información en línea; y 3) cómo se podrían usar estos datos con fines educativos. Tras el análisis de 72 webs algunos de los resultados son que solo el 24,8% incluye información detallada sobre el proyecto, y el 48,61% incluye información sobre propósitos o materiales educativos.

KEYWORDS | PALABRAS CLAVE

Citizen science, informal learning, algorithms, automatization, education, privacy protection.
Ciencia ciudadana, aprendizaje informal, algoritmos, automatización, educación, protección de la privacidad.

Received: 2022-05-30 | Reviewed: 2022-06-27 | Accepted: 2022-07-19 | OnlineFirst: 2022-10-30 | Published: 2023-01-01
DOI https://doi.org/10.3916/C74-2023-02 | Pages: 23-34

1. Introduction and state of the art

Citizen Science (CS) is the active engagement of the general public in scientific research tasks (Vohland et al., 2021). CS activities are typically organized in projects with a strong online presence via web pages and platforms which are used as data dissemination, participation and repository tools (Vohland et al., 2021). There are several international CS associations: The Citizen Science Association (CSA-North America), the European Citizen Science Association (ECSA) and the Australian Citizen Science Association (ACSA). In addition, there are national or regional associations such as *Observatorio de la ciencia ciudadana* (Spain) or Bürger schaffen Wissen (Germany) or individual projects such as Cities-Health. Information on CS activities can also be found on the websites of research institutes, universities, museums, etc. The variety of CS institutions demonstrates that communication about projects can be done through different channels (individual, as part of a 'network' or association, at local, regional or larger scale). Although the communication approach will vary throughout the project and might be different for each type of project, it is important to define it well in order to engage, retain, motivate or inform volunteers (Vohland et al., 2021; Veeckman et al., 2019). As Lin-Hunter et al. (2020) concluded in their analysis about the volunteers' tasks described in the CS project description and its connection to participant's scientific literacy development, how CS project is communicated may affect volunteers' engagement and might imply changes on public science perception and awareness of the problem to be addressed.

The Internet (through websites) or the television has historically contributed to informal science learning and science communication (Stocklmayer et al., 2010). The existence of various formations in CS demonstrates that communication about projects can be done through different channels (individual, as part of a "network" or association, at local, regional or larger scale). The materials provided on these platforms have a great potential to be used for educational purposes, especially in relation to Sustainable Development Goals (SDGs) taking into account that many CS projects address sustainability issues (Fraisl et al., 2020; Storksdieck et al. 2016). However, although multiple projects are collected in the national or global platforms, there is no centralized database that contains global information about all CS projects (Vohland et al., 2021).

Among the potential educational benefits of CS activities, we see the improvement of, scientific knowledge and understanding, the development of technical/scientific skills, STEM career motivation and values such as sustainability or respect for the environment (Hiller & Kitsantas, 2014; Bonney et al., 2016; Kobori et al., 2016; Vohland et al., 2021). Although CS projects do not usually primarily aim at fostering citizen's scientific literacy and knowledge, they often develop educational materials or conduct training activities to prepare participants for participating in scientific activities such as collecting or classifying data (Bonney et al., 2009). More and more frequently, the participation of schools in CS projects is promoted by institutions (e.g. the Oficina de Ciencia Ciudadana in Barcelona has an open call for schools to participate in CS projects: https://bit.ly/3cB1lMH), and this is increasing. However, there is still a lack of knowledge about how CS can be more centrally integrated in schools as a guide or source of inspiration for teachers to create activities aligned with current research and societal problems addressed by the CS projects. All the materials and data generated by CS projects could be used for students' learning about specific topics or support teachers' practice. This is a task for both scientists and educators to work together, so communicating science (through workshops, learning activities or informal conversations) might have an impact on the public understanding of scientific facts and knowledge (Bickford et al., 2012; Stocklmayer et al., 2010).

Given the massive presence and availability of online information on CS projects and activities, it appears promising to use computational analytics techniques to generate specific insights into the functioning and evolution of CS activities. There are many fields in which such tools have been used, especially to massively extract data from online sites and store these in databases (Diouf et al., 2019). There are few examples of use in the CS field (Ponti et al., 2018).

From a European perspective, there is a specific interest in better understanding the role of CS in science and society, e.g., the actual distribution and contribution in geographical regions, distribution over disciplines, as well as the importance of science communication in the CS field and the impact on education. There is still a lack of knowledge as to how CS projects are distributed for further developing and supporting

specific types of CS (Warin & Delaney, 2020). The work reported here is part of the EU project CS Track (https://cstrack.eu/) that operates in this line of research. For this purpose, CS Track relies on a combination of web analytics techniques and classical social studies methods. CS Track has built up a database comprising information about 4,949 CS projects that were gathered from different sites. This is the basis for the on-going extraction and further enrichment of descriptive information related to these projects. All the data centralized will allow us to know more about how CS is communicated online and to broaden our knowledge on the connections to education.

In this paper, we explain how to build a central point of knowledge about CS using as a base the information about CS projects distributed on different websites. It will allow us to see the differences and similarities between the data structures of the websites to report the data. As part of this data extraction and analysis, we have particularly tried to identify the potential for supporting educational purposes. In this work, we have been aware of constraints that are legitimately imposed by privacy and data protection principles, especially the European General Data Protection Regulation (GDPR). The GDPR aims to give citizens control over their personal data and enforces the anonymization of data unless there is no specific individual consent. A dataset can be considered anonymous if a person can't be re-identified (Gruschka et al., 2018). Although the data extracted about CS projects describes project characteristics, sometimes direct or indirect personal data is informed through the texts. The work reported here has been guided by the following research goals:

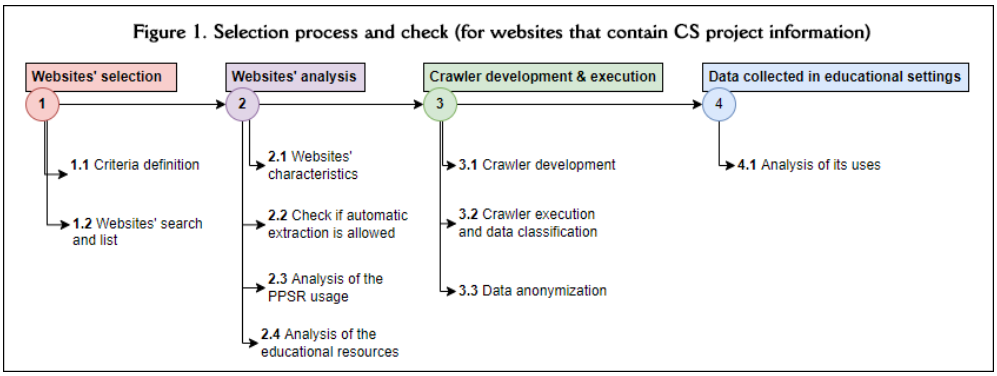
- (RG1) Design and implement an automatic algorithm to extract data from CS platforms in a unique central point (database). The extracted data should be aligned with the PPSR metadata, extended if necessary.
- (RG2) Find technical solutions to comply with GDPR requirements in this context.
- (RG3) Identify the potential educational uses of the data collected.

2. Methodology and data selection

The source of information for this study was websites that contain information about CS projects. The following criteria of inclusion was applied to identify online data from CS projects (unit of analysis):

- The website contains a list of CS projects information or are the websites of a single project.
- From Europe, associated countries or are fully conducted online.
- It is allowed to extract the data either automatically or manually.

Figure 1. Selection process and check (for websites that contain CS project information)



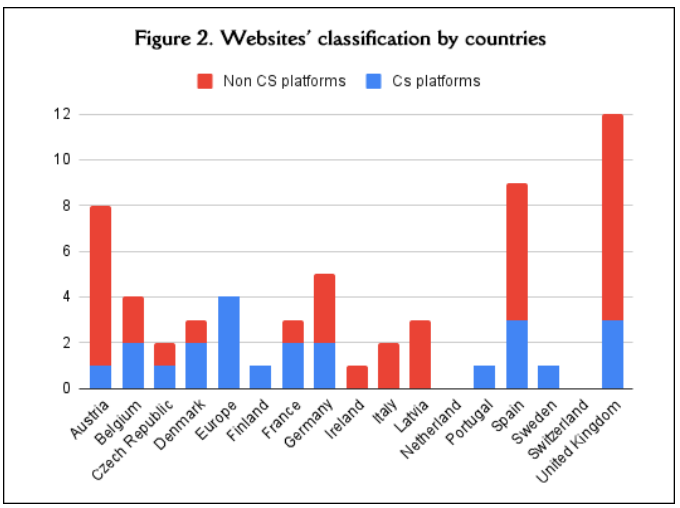
In the first phase, all the consortium members were asked to do manual online research of all the websites that could contain information about CS in European regions. After this, we manually explored each one to identify which ones contain specific information about CS projects and follow the criteria defined below. The websites' identification, selection and analysis were done manually and consist of 72 online sites. This list can be extended in next iterations. It is possible we could not identify all the existing websites that follow the criteria but, the most relevant ones were selected. The manual analysis of the website had two main objectives: (1) to identify how information of CS projects is reported, the main elements of information, the geographical distribution of websites and languages (2) to understand

the technical structure of the data and how it can be aligned with the PPSR metadata standard. Figure 1 shows the process followed during this research.

3. Citizen science presence in online platforms

This section explains the second phase of the process followed (Figure 1) and the results and findings obtained from the analysis. We classified the websites into two categories: CS platforms (29 websites) and non-citizen science platforms (43 websites). CS platforms are those digital platforms that share information about CS projects, activities, events, materials or resources, news about the field, communication tools (i.e., comments or forums) or sometimes, they are also used as a participatory tool (Sanz et al., 2019). Non-citizen science platforms' first objective is not to inform about CS, but has been created as a communication tool, as a repository or even to allow user's interaction. We analyze platform descriptions obtained from the project websites. For those classified as non-citizens' science platforms we can find a diversity of associations (i.e. Helmholtz Association), museums (i.e. Natural history museum UK) or research institutes (ICM Divulga). In the description, these websites use terminology such as "national scientific communication networking center", "Museum", "independent non-profit organization" or "Research Transfer Office" to define the association or organization.

The classification of CS platforms has been carried out following the criteria proposed by Vohland et al., (2021) which differentiates between five types of platforms. Due to the criteria followed for data selection, the category "World-wide citizen science platform" has been added for those platforms that have projects from all over the world. After analyzing the platform descriptions, we categorized them into: Commercial Platforms for CS Initiatives (2 websites), CS Platforms for Specific Projects (8 websites), CS Platforms for Specific Scientific Topics (2 websites), National CS Platforms (15 websites), EU Citizen Science Platforms (1 website) and World-wide citizen science platform (2 websites). In these texts we read the terms such as "citizen science portal" or "online citizen science hub" which are used to identify it as CS platforms and others such as "center of citizen science" or "citizen science network" in reference to the CS associations that coordinate the website. It is common for CS projects to use websites as a participatory tool, for this reason, when we read the CS platforms for specific projects, they use terms such as "simulator" or "webtool".



Europe is a continent in which cultures and languages coexist. To understand the distribution of websites across Europe, the websites have been analyzed from two points of view: the geographical location of the platform and the languages available. A total of 17 out of 44 countries have been identified in the list of websites. Figure 2 shows the countries distribution by the two types of platforms. All the online platforms considered to be "World-wide" such as SciStarter (https://scistarter.org/), iNaturalist network (https://www.inaturalist.org/), Zooniverse (https://www.zooniverse.org/) and Instant wild (https://instantwild.zsl.org/intro) have been excluded because, although we could assign to each one

a single country, they share information about projects or initiatives from all over the world. In order to better understand this geographical distribution and the citizen outreach they could achieve, it is important to also understand the linguistic diversity of Europe. Several online platforms facilitate the use of more than one language. For instance, 29.7% platforms facilitate the use of two languages (i.e. Iteritalia), 8.1% of platforms facilitate the use of three languages (i.e. OpenSystems UB) and 4.1% of platforms facilitate the use of more than three languages (i.e. EU Citizen science). 58.1% of platforms only support the use of one language (i.e. Desqbre). As stated in the Charter of fundamental rights of the European Union (European Union, 2010), "the union shall respect (...) linguistic diversity" as well as "any discrimination based on (...) language (...) shall be prohibited". The EU has 24 official languages and other regional and minority languages. 17 out of 24 languages have been covered by the websites selected. Moreover, three regional languages (Euskera or Catalan) and one extra community language (Arab) have been identified as languages in them. CS platforms cover 84.3% of languages identified.

Although we indicate which country each website can be related to, when we read the descriptions, we realized that 35.6% provides information about the region covered (such as "in Flanders" (Citizen Science Vlaanderen) or "Globally"). The Websites cover regional or national areas (i.e., Barcelona CS platform (Barcelona)), Europe (i.e., EU Citizen science platform) and all areas of the world (i.e., iNaturalist). Geographic region covered by the online platforms is aligned with the language available.

From the same information used to classify websites, key terms have been extracted to assign specific research areas. Terminology such as "protecting our planet", "science used in the investigation of crime science, laboratory analysis and the presentation of scientific evidence within the courts" or "meteorological and geophysical services" has been selected to identify the category. The platforms were classified into the six broad research areas defined in Web of Science Core Collection (Clarivate analytics, 2022): Arts & Humanities (1.37%), Life Sciences & Biomedicine (50.68%), Physical Sciences (1.37%), Social Sciences (2.74%), Technology (0%) and All (43.84%).

3.1. Websites functionalities and applications

For this research, we applied manually the platform's taxonomy defined by Derave et al. (2020). Although it defines seven categories, we have only analyzed the first three due to the websites selected being participation and communication oriented and we focus our attention on this.

Table 1. Platforms classification based on "Market side", "Affiliation" and "Centralization"

	Zero-side		**One-side**		**Multi-side**			
Market side	All websites	CS platforms	All websites	CS platforms	All websites		CS platforms	
	20.83%	17.24%	52.78%	41.38%	26.39%		41.38%	
Affiliation			**Registration**					
			All websites		**CS platforms**			
			43.06%		37.93%			
			Subscription					
			All websites		**CS platforms**			
			51.39%		55.17%			
					No transaction		**Transaction**	
					All websites	**CS platforms**	**All websites**	**CS platforms**
					79.17%	72.41%	20.83%	27.95%
							Investment	
							0%	
Centralization					**Centralized**			
					All websites		**CS platforms**	
					33.33%		51.72%	
					Decentralized			
					All websites		**CS platforms**	
					66.67%		48.28%	

- Market sides: It is the first category that defines the number of user groups. We also included the term Zero-side. From the selected sites, we identified 15 sites Zero-side (no interactions between users, only between them and website manager), 38 One-side (users' interaction is

with the platforms and indirectly with other users, via comments or posts) and 19 Multi-side (interaction within the users and platform). CS platforms are commonly One-side or Multi-side oriented although, as they are sometimes used as a participatory tool, it is common that they allow users to interact with each other (i.e., via forums or comments) (Table 1).

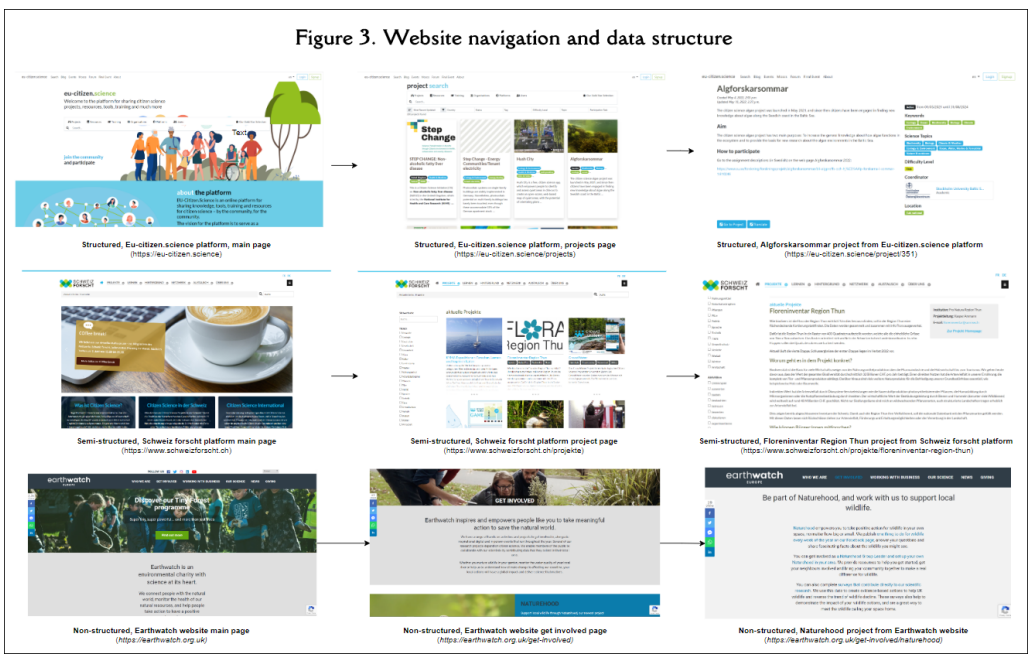
- Affiliation: It refers to how the users interact with others and the website. There are many functionalities designed for giving users the opportunity to be connected (i.e., forums, newsletters, etc.) and can be combined in a single site. We identify 31 sites that allow registration, 37 that allow subscription, 19 that allow main content creation or commenting, 15 that allow transaction and 0 that allow investment. For Multi-side platforms it is common to involve users in commenting or content creation, ask for registration and give the option of being connected and informed via subscription. Nevertheless, for the other types, registration is not a main requirement but subscription is highly recommended to be connected. The most common tool used for user interaction is the forum, but only 8 websites have one (5 of them are CS platforms). The second option for user interaction is adding comments (no direct messages). Only 6 of the platforms allow this type of functionality. As sites for scientific research, dissemination of results and participant" engagement is important. There are special pages for news or a blog created in these sites: 47 of them have one, 20 of them being for CS platforms.
- Centralization: the third category is aligned with the second one because it indicates the way the users connect among each other. There are 48 Decentralized websites while there are 13 Centralized websites.

3.2. Citizen science information online

In this section, it is explained how CS information is shared in websites from two points of views: How data are structured and what kind of data are shared. The analysis is necessary for the algorithm and database design.

Although each web page follows its own design and data structure, it is common for all to have a main page, then a page with the list of projects and indexed links to another page with the information of the associated CS project (Figure 3). However, there is an exception for CS platforms for specific projects since it contains information about a single project distributed in pages, not a list of individual projects as the others.

Figure 3. Website navigation and data structure



As a first step, the original data source was classified according to three types of web pages (Figure 3):

- Structured: information is presented in categories (i.e. https://eu-citizen.science/).
- Semi-structured: some information is presented incategories and another is presented in paragraphs (i.e. https://www.zooniverse.org/).
- Non-structured: no information categorized (i.e. https://www.scivil.be/en).

From the 72 websites we can identify 13 websites with structured data, 27 with semi-structured data and 34 with non-structured data. Several working groups from CS associations (Data & Metadata working group (CSA) or Working Group on "Data, Tools and Technology" (ECSA)) are focused on promoting standardization of CS data. This is the case, for example, of the Public Participation in Scientific Research - Core (PPSR-Core) data model which proposes a data standard and works on promoting it to be accepted and used by CS websites (Bowser et al., 2017). Regarding the CS projects information, in order to analyze how this data is shared online and define a common structure for the database to store the data classified, it has been necessary to identify how the information associated with a CS project can be classified according to the PPRS-Core metadata standard attributes and some that are newly created. We analyzed how this standard is followed in the websites, and the Title was well identified (commonly it is the first shown and bigger than other texts) and Description (below the title). For the other categories, we have created a dictionary of terms containing 19 categories (15 included in the metadata standard and 4 added to the standard), based on similar terms/sections contained in the different websites analyzed:

- Social media: the name of the social media platform (i.e., "twitter" or "facebook") or general terms (i.e., "blog.", "REDES SOCIALES:" or "PERFILES EN REDES SOCIALES:").
- Online resources: file extension formats (i.e., ".pdf) or general terms (i.e., "OTROS RECURSOS DEL PROYECTO:", "Desktop:").
- Tools and materials: only one expression selected "Tipo de medios".
- Applications used: applications repositories names although could be integrated into Tools and materials category (i.e., "play.google" or "apple.com") or general terms (i.e., "Mobile:").
- Images: images file extension (i.e., ".jpg", ".png", ".JPG" or ".jpeg").
- Geographical location: general terms used (i.e., "Geographical", "Geographic Scope", "WHERE", "Ubicación", "places", "Project Location" or "Location") or specific terms for regions or areas (i.e., "Country", "PROVINCIA:" or "País ").
- Status: general terms (i.e., "Project Status", "Status" or "ESTADO DEL PROYECTO:").
- Methodology - Participants tasks: general terms (i.e., "Participation Tasks" or "Tasks") and open questions about the participation (i.e., "HOW TO GET STARTED", "RELACIÓN CON LA CIENCIA CIUDADANA:" or "¿Cómo participan los voluntarios/as?").
- Start date: general terms (i.e, "Start Date", "FECHA DE INICIO DEL PROYECTO:" or "Projektstart:").
- Investment or support: general terms (i.e., "Sponsor", "TOTAL EXPENSE" or "Project Funding").
- Field of science: general term (i.e., "Fields of Science", "TOPICS", "ÁREA DE CONOCIMIENTO:").
- Development time: general terms (i.e., "Intended Outcomes", "IDEAL FREQUENCY", "When? " or "Période : ").
- Main objectives: general terms (i.e., "Goal", "Waarom doe je mee?" or "Objet : ").
- Participants age: only the term "IDEAL AGE GROUP".
- Participants profile: general terms (i.e., "Wie kan meedoen?", "Usuarios", "/people/", "PÚBLICO AL QUE SE DIRIGE EL PROYECTO:", "INTEGRANTES DEL PROYECTO:", "INTEGRANTES:", "Who can take part?", "Public: ", "Project Partners" or "Users").
- Development place: general terms to explain the space or area to develop activities research (i.e., "SPEND THE TIME", "Region", "Ubicación", "ÁMBITO DE ACTUACIÓN:" or "Type of activity:").
- Dedication time: general terms to explain how much time participants will invest in participation (i.e., "AVERAGE TIME" or "How long will it take? ").

- Contact information: "@" is used in the email addresses.
- Project update date: only the term "PROJECT UPDATED".
- "Main program or person in charge": in this category it is combined information about the information creator, coordinators or managers and associations that support or collaborate (i.e., "PRESENTED BY", "Wie organiseert het?", "/researcher/", "Creado por:", "Administradores de proyecto:", "Administrador de proyecto:", "MIEMBROS DEL EQUIPO:", "OTROS GRUPOS O INSTITUCIONES COLABORADORES:", "OTRAS PERSONAS O ENTIDADES COLABORADORAS:", "Project Manager", "Project Co-ordinator" or "Kontakt:").

4. Algorithm development and execution

Our goal is to have all the information stored in a database, so it is essential to choose the correct ones based on the type of data extracted from the online platforms selected. In this section, the process followed in the third phase of the process is described (see section 2).

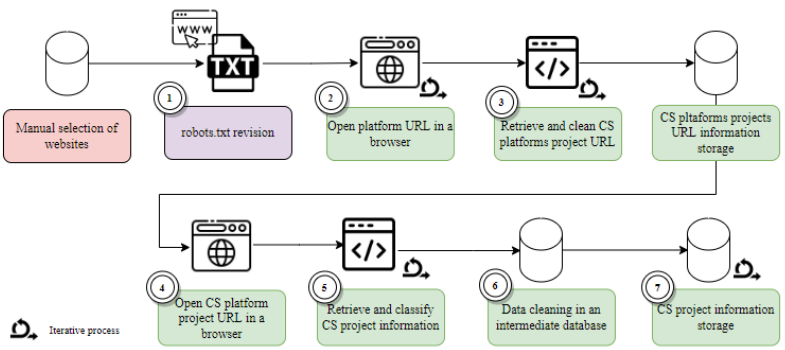
4.1. Database selection

In order to select the database, the comparison of the type of databases was made between relational databases (those accepting Structured Query Language, (SQL)) and non-relational databases (those which do not accept SQL) (Li & Manoharan, 2013). MongoDB database version 4.2 was selected because it can store structured and unstructured data; it can easily grow; the database structure can be changed independently to other data collections and documents (data structure depends on each project description) and it admits queries and data consumption.

4.2. Algorithm development and execution

The algorithm is adapted to the three types of web structures. In accessing CS project information, we applied a two-step process: first to access a main page where projects are listed, and second, select a certain CS project to see its information. Figure 4 shows the process followed by the algorithm. It was developed in Python programming language (using selenium, b4soup, requests and PyMongo libraries). In this process, it was necessary to take into account if the web pages had an Application Programming Interface (API) (the EU. Citizen science and iNaturalist websites) which allows the automatic extraction of data from the database source.

Figure 4. Process followed by the algorithm



The robot exclusion protocol regulates, for bots, access to the source code of the website (Kolay et al., 2008). It is defined by each website and informed in the robots.txt file which is accessible via the website URL. The restrictions combine blocking all or partial content for all or certain bots. This information is checked before data extraction. We identified: 4 websites didn't define robots' exclusion protocol; 20 websites contain restrictions but not for the algorithm created or the specific content to extract and 1 website whose protocol does not allow ".pdf" extraction. For each website, the parser searches the source code manually identified previously and extracts specific web elements (Parvez et al., 2018). Furthermore,

user actions such as pressing buttons have to be replied to because some pages which contain CS project information have more than one tab.

4.3. Data classification, cleaning and storing

The CS project data classification is done automatically searching by keywords, symbols or sentences previously defined (See section 3.3). A cleaning process was applied to remove wrong data or all the elements unclassified and the final information is stored in the database. To avoid duplicates, it is checked if the project title exists in the database. Additional information such as storage date, update date (in case it already exists), origin or ID is added before storing. The information obtained through the API has to be mapped to the defined database structure (cleaning is not needed). This process can be reproduced when needed, so the number of projects and information can increase. Figure 5 shows an example of the information stored.

Figure 5. The Neureka project information stored in the database


```
{
    "_id": {
        "$oid": "5efc3ae868516f3f3e568278"
    },
    "TITLE": "The Neureka Project",
    "WEB": ["https://www.neureka.ie/", "WEBSITE The Neureka Project"],
    "SOCIAL MEDIA": ["https://www.youtube.com/watch?v=TkWtYafAAT4", "https://twitter.com/@neurekaApp"],
    "DESCRIPTION": ["Neureka is a collection of fun brain games and challenges that allows you to solve cogr
    "Plat Id": ["65", "17"],
    "Plat country": ["USA", "EU"],
    "Insert date": "2020-07-01",
    "APPS": ["https://play.google.com/store/apps/details?id=com.gillanlab.neureka.beta", "https://apps.apple
    "GEOGRAPHICAL LOCATION": ["WHERE Online"],
    "METHODOLOGY": ["HOW TO GET STARTED Download neureka from either the Google Play Store or Apple App Stor
    "INVESTMENT": ["TOTAL EXPENSE 0.00"],
    "TOPICS": ["Biology", "Psychology", "Health & Medicine", "brain", "dementia", "depression", "neuroscienc
    "DEVELOPMENT TIME": ["IDEAL FREQUENCY Twice per day"],
    "DEDICATION TIME": ["AVERAGE TIME 1 hour"],
    "DEVELOPMENT SPACE": ["SPEND THE TIME Indoors"],
    "PLATFORM UPDATE DATE": ["PROJECT UPDATED 06/22/2020, 08:25 pm GMT+2", "2021-03-03T10:35:12.479221Z"],
    "Url platform": ["https://scistarter.org/the-neureka-project", "https://eu-citizen.science/projects"],
    "MAIN PROGRAM OR PERSON IN CHARGE": ["PRESENTED BY Gillan Lab at Trinity College Dublin", "Trinity Colle
    "MAIN OBJECTIVES": ["GOAL Improve mental health and dementia research"],
    "TOOLS AND MATERIALS": ["MATERIALS A smartphone (either Android or Apple)"],
    "PARTICIPANTS PROFILE": ["SPECIAL SKILLS None"],
    "Wp2 Id": ["3", "17"],
    "Language": ["English"],
    "STATUS": ["Active"],
    "START DATE": ["2020-07-01"],
    "END DATE": ["2025-12-31"],
    "LATITUDE": ["53.343667"],
    "LONGITUDE": ["-6.254445"],
    "COUNTRY": ["IE"],
    "IMAGE": ["https://eu-citizen.science/media/media/images/2021-02-08_020202020404_546_NEUREKA%20LOGO.png"
    "Date update": ["2021-03-19", "2021-04-08", "2021-05-14"],
    "MAIL": null
}
```


We use Named Entity Recognition (NER) paired with the Entity Ruler to identify phone numbers, email addresses and personal accounts based on given regular expression (RegEx) patterns. The algorithm then checks if names of individuals are occurring in connection with personal data found by the Entity Ruler. Names can in some cases remain not-anonymized, such as if there is a wikipedia article for that name, as this indicates that this is either a person in the public eye and their name carries meaning beyond naming a person (i.e. Albert Einstein), or it is a common name that does not identify one single person. Another reason can be if the person's name has an overlap with the project name as we found that some projects center around analyzing data about a single person or a family. Both these reasons are disregarded however if a phone number, email or personal account information is mentioned in the same sentence.

For the anonymization we opted to go for a pseudonymization by repeatedly hashing the selected text parts (Kasgar et al., 2012) and adding the previously ascribed type of data into the text. This way the individual persons as well as the personal data cannot be identified, but we retain some readability of the text and connections over multiple projects via common persons are still possible.

5. Results

In our analysis, we included 4 new categories in addition to the existing 43 categories included in the PPSR-Core standard: 11 required fields, 21 optional fields and 15 new attributes identified in the websites. This classification is aimed to facilitate the automatic analysis of data in next steps. The main

problem identified is the inconsistency when reporting data about CS projects online, since most of the websites or platforms do not follow metadata standards. The most common attributes (required) are title, description, website link (if exists), social media, contact information or project topic. 91.56% of them are informed. All the information related to data origin, language, storage date and other information necessary for data management is always informed but is not considered for this analysis. Less common attributes (defined as optional or news created by us) such as geolocation or member's age are informed in 24.8%. Data mining techniques have been used to extract data from texts in other fields. An important source of information is the project description, these techniques can be applied to automatically extract information from this category and fill other attributes in the database. For instance, as explained in previous sections, NER has been used to identify persons' names but it can also be used to get information from the text about the organizations, cardinal numbers, countries, cities or states, etc.

Other computational methods oriented for data classification can also be applied to the data in order to add categories or create clusters to easily target the data for easy filtering or give them a context (Roldán-Álvarez et al., 2021). By extracting keywords or meaning from the texts, new categories such as: sustainable development goals (SDGs), learning outcomes or research areas can be created. Result of the methods application can be used to support teachers to select the topic and get inspiration to create learning activities.

6. Discussion and conclusions

The variety of websites that share information about CS projects is also a reflection of how variable CS is. Involving citizens also implies adapting to different forms of communication, either because of the language or the region in which it occurs. Websites are great tools for this communication and sharing with others but also for participation. Even so, there is still work to be done in order to increase public access for CS to be well known and to increase citizen's interest in participating in research. CS platforms should consider being aligned with the PPSR_Core and other metadata standards. Normalizing all the data structures and information shared improves the user's experience in the websites along with facilitating them the search. Having the key information about a CS project all together and well documented could also improve citizen's participation and interest and the research analysis of the CS field. In this line, CS Track projects, besides developing knowledge on the CS field, had opened a new perspective on how computational methods can be applied to centralize all the data into a single database for research purposes. There is still work to be done to analyze and apply data mining methods to the data in order to obtain more information for the empty categories. Nevertheless, the text mining methods are useless without a good and detailed CS project description. It is necessary to involve scientists and communication experts (Roche et al., 2020) and follow guidelines already defined by experts (as the one proposed by Veeckman et al., 2019) for good communication action. In order to have educational impact, it is essential to be aligned with the official curriculum of the educational level to which they refer.

6.1. How to identify content with potential educational benefit?

CS platforms and websites can provide content that can be used as a powerful resource for learning and teaching. A first exploratory study developed by Calvera-Isabal et al. (2021) has explored three CS platforms and found that materials and data related to CS projects extracted from websites have the potential to support teachers in their practice (Asensio-Pérez et al., 2014). Previous publications have stated that data exploration has an impact on student's awareness and interest and promotes discussion, opening new perspectives on how to work mathematics in formal education (Saddiqa et al., 2019). From the data classified into categories, teachers will find a powerful source of scientific knowledge for filtering (for instance, based on Research Areas, SDGs or learning skills). Classifying the data by research areas will allow teachers to better understand the field the project is investigating. Information from SDGs, which are addressing world-wide real problems, can be integrated into the learning designs to motivate students to learn more about and also create awareness about, for instance, sustainability or ecology (Massa et al., 2011; Djonko-Moore et al., 2018). Although CS is being integrated in education and has potential to be integrated in many other ways, only 48.61% contain educational material or information related to learning.

For CS platforms there is a more positive result as 55.17% have these resources. The ones that allow citizens participation have specific pages with educational materials and common questions answered (Zooniverse). Specific CS project platforms are used for communication but also as a repository tool for all the information and documentation they develop (Luonto-Liiton Kevätseuranta). The other websites that are not CS platforms, if they are education oriented (i.e., universities or associations about education or learning), then they share specific materials for teachers or educators. If not, it is not common to share these types of resources. The application of advanced computational techniques and having all the information centralized, can be used to support information online: real problems, research areas, scientific disciplines, learning skills, etc.

Finally, it is common for teachers to integrate technology to support learning or enhance it. For this reason, tools and content developed by CS projects might be integrated in the classroom as an instrument to develop an activity, to participate in CS or even to support them during the lesson preparation. Regarding the potential usage of the data in educational contexts covered in this article, it is also necessary to work more on identifying how to communicate (at the level of data/information to be reported on CS websites) to narrow the link that may exist between CS and formal educational contexts. Some opportunities that arise from this analysis are the usage of the CS project information in educational contexts (such as to inspire teachers to create learning design activities) or the participation of schools in the project (such as particular follow-up cases). It is still necessary to analyze the materials teachers need and to what extent all this information and resources supports them for their teaching practice. It is expected that all this data and resources centralized and available to be explored, have an impact on teacher's scientific knowledge and pedagogical skills, which might affect student's attitude toward science (Chan & Yung, 2018). Finally, the application of algorithms and the collection of mass information allows the unification of data in a single source that could potentially be used for educational purposes. For this, as future work, a digital platform could be developed that communicates CS information to support the creation of activities in the classroom.

Authors' Contribution

Idea, M.C, P.S, U.H; Literature review (state of the art), M.C, P.S, U.H, C.S; Methodology, M.C, P.S, U.H; Data analysis, M.C, C.S; Results, M.C, P.S; Discussion and conclusions, M.C, P.S, U.H, C.S; Writing (final draft), M.C, P.S, U.H, C.S; Final revisions, P.S., U.H; Project design and funding agency, M.C, P.S, U.H.

Funding Agency

This work was supported in part by PID2020-112584RB-C33 funded by MCIN/AEI/10.13039/501100011033, by 'CS Track: Expanding our knowledge on Citizen Science through analytics and analysis' H2020-SwafS-2019-1 proj. ref: 872522 and by the Ramón y Cajal programme (P. Santos).

References

Asensio-Pérez, J.I., Dimitriadis, Y., Prieto, L.P., Hernández-Leo, D., & Mor, Y. (2014). From idea to VLE in half a day: METIS approach and tools for learning co-design. In *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 741-745). ACM. https://doi.org/10.1145/2669711.2669983

Bickford, D., Posa, M.R.C., Qie, L., Campos-Arceiz, A., & Kudavidanage, E.P. (2012). Science communication for biodiversity conservation. *Biological Conservation*, *151*(1), 74-76. https://doi.org/10.1016/j.biocon.2011.12.016

Bonney, R., Cooper, C.B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K.V., & Shirk, J. (2009). Citizen science: A developing tool for expanding science knowledge and scientific literacy. *BioScience*, *59*(11), 977-984. https://doi.org/10.1525/bio.2009.59.11.9

Bonney, R., Phillips, T.B., Ballard, H.L., & Enck, J.W. (2016). Can citizen science enhance public understanding of science? *Public Understanding of Science*, *25*(1), 2-16. https://doi.org/10.1177/0963662515607406

Bowser, A., Brenton, P., Stevenson, R., Newman, G., Schade, S., Bastin, L., Parker, A., & Oliver, J. (2017). *Citizen Science Association Data & Metadata Working Group: Report from CSA 2017 and Future Outlook*. European Commision. https://bit.ly/3IS85Sl

Calvera-Isabal, M., Varas, N., & Santos, P. (2021). Computational techniques for data science applied to broaden the knowledge between citizen science and education. In D. G. Sampson, D. Ifenthaler, & P. Isaías (Eds.), *Proceedings of the 18th International Conference on Cognition and Exploratory Learning in the Digital Age (CELDA 2021)* (pp. 219-226). IADIS Press. https://doi.org/10.1007/978-3-030-65657-7

Chan, K.K.H., & Yung, B.H.W. (2018). Developing pedagogical content knowledge for teaching a new topic: More than teaching experience and subject matter knowledge. *Research in Science Education*, *48*, 233-265.

https://doi.org/10.1007/s11165-016-9567-1

Clarivate analytics (Ed.) (2022). *Web of Science Core Collection Help*. https://bit.ly/3ts2IZI

Derave, T., Sales, T.P., Gailly, F., & Poels, G. (2020). Towards a reference ontology for digital platforms. In G. Dobbie, U. Frank, G. Kappel, S. W. Liddle, & H. C. May (Eds.), *Conceptual modeling. ER 2020. Lecture notes in computer science* (pp. 289-302). Springer. https://doi.org/10.1007/978-3-030-62522-1_21

Diouf, R., Sarr, E.N., Sall, O., Birregah, B., Bousso, M., & Mbaye, S.N. (2019). Web Scraping: State-of-the-Art and Areas of Application. In *2019 IEEE International Conference on Big Data (Big Data)* (pp. 6040-6042). https://doi.org/10.1109/BigData47090.2019.9005594

Djonko-Moore, C.M., Leonard, J., Holifield, Q., Bailey, E.B., & Almughyirah, S.M. (2018). Using culturally relevant experiential education to enhance urban children's knowledge and engagement in science. *Journal of Experiential Education*, *41*(2), 137-153. https://doi.org/10.1177/1053825917742164

European Union (Ed.) (2010). *Charter of fundamental rights of the European Union. Official Journal of the European Union C83*, *53*, *380*. https://bit.ly/3PFo605

Fraisl, D., Campbell, J., See, L., Wehn, U., Wardlaw, J., Gold, M., Moorthy, I., Arias, R., Piera, J., Oliver, J.L., Masó, J., Penker, M., & Fritz, S. (2020). Mapping citizen science contributions to the UN sustainable development goals. *Sustainability Science*, *15*(6), 1735-1751. https://doi.org/10.1007/s11625-020-00833-7

Gruschka, N., Mavroeidis, V., Vishi, K., & Jensen, M. (2018). Privacy issues and data protection in big data: A case study analysis under GDPR. In *2018 IEEE International Conference on Big Data (Big Data)* (pp. 5027-5033). IEEE. https://doi.org/10.1109/BigData.2018.8622621

Hiller, S.E., & Kitsantas, A. (2014). The effect of a horseshoe crab citizen science program on middle school student science performance and STEM career motivation. *School Science and Mathematics*, *114*(6), 302-311. https://doi.org/10.1111/ssm.12081

Kasgar, A.K., Agrawal, J., & Sahu, S. (2012). New modified 256-bi MD5 algorithm with SHA. Compression Function. *International Journal of Computer Applications*, *42*(12), 47-51. https://doi.org/10.5120/5748-7956

Kobori, H., Dickinson, J.L., Washitani, I., Sakurai, R., Amano, T., Komatsu, N., Kitamura, W., Takagawa, S., Koyama, K., Ogawara, T., & Miller-Rushing, A.J. (2016). Citizen science: a new approach to advance ecology, education, and conservation. *Ecological Research*, *31*(1), 1-19. https://doi.org/10.1007/s11284-015-1314-y

Kolay, S., D'Alberto, P., Dasdan, A., & Bhattacharjee, A. (2008). A larger scale study of robots. txt. In *Proceedings of the 17th international conference on World Wide Web* (pp. 1171-1172). https://doi.org/10.1145/1367497.1367711

Li, Y., & Manoharan, S. (2013). A performance comparison of SQL and NoSQL databases. In *2013 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)* (pp. 15-19). IEEE. https://doi.org/10.1109/PACRIM.2013.6625441

Lin-Hunter, D.E., Newman, G.J., & Balgopal, M.M. (2020). Citizen scientist or citizen technician: A case study of communication on one citizen science platform. *Citizen Science: Theory and Practice*, *5*(1). https://doi.org/10.5334/cstp.261

Massa, N., Dischino, M., Donnelly, J.F., & Hanes, F.D. (2011). Creating real-world problem-based learning challenges in sustainable technologies to increase the STEM Pipeline. In *2011 ASEE Annual Conference & Exposition*. https://doi.org/10.18260/1-2–17678

Parvez, M.S., Tasneem, K.S.A., Rajendra, S.S., & Bodke, K.R. (2018). Analysis of different web data extraction techniques. In *2018 International Conference on Smart City and Emerging Technology (ICSCET)* (pp. 1-7). IEEE. https://doi.org/10.1109/ICSCET.2018.8537333

Ponti, M., Hillman, T., Kullenberg, C., & Kasperowski, D. (2018). Getting it right or being top rank: Games in citizen science. *Citizen Science: Theory and Practice*, *3*(1). https://doi.org/10.5334/cstp.101

Roche, J., Bell, L., Galvão, C., Golumbic, Y.N., Kloetzer, L., Knoben, N., Laakso, M., Lorke, J., Mannion, G., Massetti, L., Mauchline, A., Pata, K., Ruck, A., Taraba, P., & Winter, S. (2020). Citizen science, education, and learning: challenges and opportunities. *Frontiers in Sociology*, *5*, 613814-613814. https://doi.org/10.3389/fsoc.2020.613814

Roldán-Álvarez, D., Martínez-Martínez, F., & Martín, E. (2021). Citizen science and open learning: A Twitter perspective. In *2021 International Conference on Advanced Learning Technologies (ICALT)* (pp. 6-8). IEEE. https://doi.org/10.1109/ICALT52272.2021.00009

Saddiqa, M., Larsen, B., Magnussen, R., Rasmussen, L.L., & Pedersen, J.M. (2019). Open data visualization in Danish schools: A case study. *Complex Systems Informatics and Modeling Quarterly*, *21*, 1-20. https://doi.org/10.24132/CSRN.2019.2902.2.3

Sanz, F., Gold, M., & Mazzonetto, M. (2019). *D2.3: Platform functionality requirements & specification report*. Zenodo. https://doi.org/10.5281/zenodo.3612808

Stocklmayer, S.M., Rennie, L.J., & Gilbert, J.K. (2010). The roles of the formal and informal sectors in the provision of effective science education. *Studies in Science Education*, *46*(1), 1-44. https://doi.org/10.1080/03057260903562284

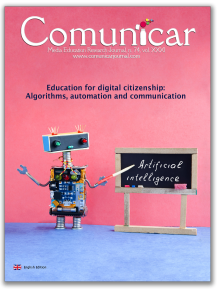
Storksdieck, M., Shirk, J.L., Cappadonna, J.L., Domroese, M., Göbel, C., Haklay, M., Miller-Rushing, A.J., Roetman, P., Sbrocchi, C., & Vohland, K. (2016). Associations for citizen science: Regional knowledge, global collaboration. *Citizen Science: Theory and Practice*, (2), 1-1. https://doi.org/10.5334/cstp.55

Veeckman, C.M., Talboom, S., Gijsel, L., Devoghel, H., & Duerinckx, A. (2019). *Communicatie bij burgerwetenschap: Een praktische handleiding voor communicatie en betrokkenheid bij citizen science*. SCIVIL. https://bit.ly/3PKQz50

Vohland, K., Land-Zandstra, A., Ceccaroni, L., Lemmens, R., Perelló, J., Ponti, M., & Samson, R. (2021). *The science of citizen science*. Springer Nature. https://doi.org/10.1007/978-3-030-58278-4

Warin, C., & Delaney, N. (2020). *Citizen science and citizen engagement. Achievements in Horizon 2020 and recommendations on the way forward*. European Commision. https://doi.org/10.2777/05286

Comunicar, n. 74, v. XXXI, 2023 | Media Education Research Journal | ISSN: 1134-3478; e-ISSN: 1988-3478
www.comunicarjournal.com



Reflections on the ethics, potential, and challenges of artificial intelligence in the framework of quality education (SDG4)

Reflexiones sobre la ética, potencialidades y retos de la Inteligencia Artificial en el marco de la Educación de Calidad (ODS4)

Dr. Jesús-Miguel Flores-Vivar. Professor, Department of Journalism and New Media, Complutense University of Madrid, Madrid (Spain) (jmflores@ucm.es) (https://orcid.org/0000-0003-1849-5315)
Dr. Francisco-José García-Peñalvo. Full Professor, Department of Computer Science and Automation, Universidad de Salamanca, Salamanca (Spain) (fgarcia@usal.es) (https://orcid.org/0000-0001-9987-5584)

ABSTRACT

This article analyses and reflects on the ethical aspects of using artificial intelligence (AI) systems in educational contexts. On the one hand, the impact of AI in the field of education is addressed from the perspective of the Sustainable Development Goals (specifically, SDG4) of the UNESCO 2030 Agenda, describing the opportunities for its use by teachers and students. On the other hand, there is an emphasis on the uncertainties caused by the fears of some who think, for example, that AI robots will replace human teachers. The methodology used is based on a documentary review, supported by reports and studies conducted by researchers as well as institutions and organizations committed to the development of AI and its capacity for action in the educational field, and the ethical questions that experts and organizations around the world are raising. The results seek to influence and deepen the ethical implications that AI can have in the educational field. Finally, some critical questions about the ethics and challenges of AI in education are proposed, concluding with the need to add proposals with new research and political actions which could lead to the creation of an ethical observatory of AI for education.

RESUMEN

El artículo analiza y reflexiona sobre aspectos coyunturales de la ética en el uso de los sistemas de Inteligencia Artificial (IA) en contextos educativos. Se aborda, por un lado, el impacto de la IA en el campo de la Educación, desde la perspectiva de los Objetivos de Desarrollo Sostenible (concretamente, ODS4) de la Agenda 2030 de la UNESCO, describiendo las oportunidades en su uso por parte de docentes y estudiantes. Por otro, se hace hincapié en las incertidumbres, sobrevenidas por los temores de algunos que piensan, por ejemplo, que los robots de IA reemplazarán a los profesores humanos. La metodología utilizada se basa en la revisión documental que parte de informes y estudios de investigadores, así como de instituciones y organismos comprometidos con el desarrollo de la Inteligencia Artificial y de su capacidad de acción en el ámbito educativo, para, a partir de ahí, abordar las cuestiones éticas que se vienen planteando por parte de expertos y de organizaciones de todo el mundo. Los resultados obtenidos buscan incidir y profundizar en las implicaciones éticas que la Inteligencia Artificial puede tener en el ámbito educativo. Finalmente, se proponen algunas cuestiones clave de la ética y los retos de la IA en la educación concluyéndose en la necesidad de sumar propuestas con nuevas investigaciones y acciones políticas, retos que podrían materializarse con la creación de un observatorio ético de IA para la educación.

KEYWORDS | PALABRAS CLAVE

Digital literacy, education, teacher training, artificial intelligence, ethics, Internet.
Alfabetización digital, educación, formación del profesorado, inteligencia artificial, ética, Internet.

Received: 2022-05-30 | Reviewed: 2022-06-28 | Accepted: 2022-07-13 | OnlineFirst: 2022-10-30 | Published: 2023-01-01
DOI https://doi.org/10.3916/C74-2023-03 | Pages: 35-44

1. Introduction, approach, and methodology

The technological evolution of recent years has had a positive and/or negative impact on societies worldwide, affecting people's modus vivendi at work, in health, economy, and, obviously, education and training (Alonso-de-Castro & García-Peñalvo, 2022).

Sustainable Development Goal 4 (SDG4) of the 2030 Agenda (United Nations, 2019), promoted by UNESCO, is also a common goal of states and governments. SDG4 seeks to "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all" (UNESCO, 2019). Furthermore, technologies play a crucial role in achieving this goal. Thus, over the last two decades, emerging technologies have disrupted teaching-learning processes at different levels of education. In this context, artificial intelligence (AI) is a potential technological tool (Vitanza et al., 2019) that, without being new, is expanding in all professional and knowledge fields, affecting, impacting, and causing a real revolution in the field of education. This branch of computer science is responsible for studying and assessing the possibilities for a machine to perform human tasks (Túñez-López & Tejedor-Calvo, 2019). According to the UNESCO (2021b) document "Recommendation on Artificial Intelligence", AI systems are information processing technologies that integrate models and algorithms that produce a capacity to learn and perform cognitive tasks, leading to results such as prediction and decision-making in material and virtual environments.

Thus, SDG4 emphasizes that AI technologies are used to ensure equitable and inclusive access to education (UNESCO, 2019). This implies having to change the role of teachers in the transmission of the knowledge they provide to the younger generations. Other authors (Osetskyi et al., 2019) argue that the future of higher education is inextricably linked to the development and corresponding increase in the capacity of new intelligent machines capable of operating large amounts of information, self-learning, and improvement, whereby AI has become a new focus of international competition of countries in the educational market. Therefore, we are witnessing how AI's presence in the educational field is constantly growing through robotics or algorithms. In fact, its immersion in this field has occurred "silently" hand in hand with other technologies already consolidated in the educational field, such as virtual campuses and academic social networks (Flores-Vivar & Zaharía, 2019). Along these lines, a series of research studies and analyses have been carried out by specialists who, through publications, seminars, and congresses, paint a good picture, but also one of the ethical solid, and deontological implications in the use of AI in the field of education. An example is the "International Forum on AI and the Futures of Education Developing Competencies for the AI Era" (UNESCO, 2021a), held in December 2020. Building on the Beijing Consensus, the forum shared policies and practices regarding the role of AI in education, with a specific focus on defining the competencies required in the AI era and strategies to prepare all people to live and work with AI effectively. The European Commission (2020) also proposes an "EU strategic framework based on EU values that will give citizens the confidence to accept AI-based solutions". For its part, the European Parliament (2021) has been legislating on the implementation and use of AI technologies in the education sector, warning that they are high-risk technologies and should therefore be subject to stricter security, transparency, fairness, and accountability requirements. Being a particular area of study in engineering and computer science, we see that AI is increasingly being applied in the field of education, supporting the management, learning, and evaluation of students. One example is the latest developments in Google Classroom (Kiecza, 2022), which uses AI to help students and teachers. Other examples include using AI to automate routine administrative tasks, diagnose student competencies, and provide learning content and feedback tailored to individual progress. Nevertheless, the implementation and development of AI technologies applied to education must be seen as a complex and highly controversial issue (Selwyn, 2019), especially regarding the ethical and deontological values that must be impregnated from the design and subsequent development phase. Hence, the importance of AI in the field of teaching–learning is growing more and more. AI is increasingly pervasive and has transcended national borders, so "all countries must work together to ensure that AI remains under human control and is designed and applied for the common good" (UNESCO, 2021a: 5). In particular, it is essential that humans are protected from becoming victims of AI tools, so we must understand that AI must be used to augment and amplify human capabilities but not replace them. Moreover, this understanding begins in education.

According to the report "Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development" (UNESCO, 2019), "Artificial intelligence can be a great opportunity to accelerate the achievement of the SDGs. However, technological revolution leads to new imbalances that we must anticipate". These imbalances are caused by the impact generated. Hence, rapid technological advances in AI and other advanced technologies, such as robotics, cloud computing, and the Internet of Things, are transforming disciplines, economies, and industries and challenging ideas about what it means to be human.

AI has enormous potential for social good and for promoting the achievement of SDGs if it is developed in a way that benefits humanity, respects global norms and standards, and is anchored in peace and development. Considering this, this paper aims to explore the contexts, uses, and conditions that favor –or not– the application of AI, along with the ethical and regulatory disjunctions that may be critical in the field of education. The background is that the ethical dilemma related to the understanding and interpretability of AI agents' behavior is one of the fundamental challenges of the next decade in this area (Flogie & Aberšek, 2021).

On the one hand, this study seeks to analyze and explain an existing problem, which has been addressed for some years according to the scientific literature reviewed, the rise of AI in education, the unknown about its uses and ethical aspects in its development, and factors that affect the users of these technologies (teachers and students). On the other hand, it proposes new conjectures, challenges, and paradigms that, together with the implementation and consolidation of AI in education, should be considered for future – and present – generations of these groups.

Therefore, the methodology used in this work is based on a documentary review of primary and secondary sources on the development of AI and its capacity for action in the educational field, consistent with SDG4 for quality and inclusive education. This review has been done considering that the literature on the subject is still scarce, "focusing more on the computer part and less on the impacts of AI on society" (Canavilhas, 2022). In the criteria for the selection of sources from international institutions and organizations, such as UNESCO, United Nations, OECD, European Commission, or European Parliament, it was considered that they are organizations that seek to strengthen the development of quality and inclusive education worldwide and that, for years, have been addressing the use of AI in this area. As for the scientific papers analyzed, the criteria were based on the specific subject matter, the level of the publications, and the expert positions represented by many authors, considering that it was not possible to include all of them for space reasons. From these analyses, we seek to extrapolate and address the ethical and philosophical issues that have been raised by experts and organizations around the world, seeking reflection, debate, and foresight on the profound ethical and deontological implications that AI, in general, will have in the global educational field. According to UNESCO (2021a), regulation is necessary. With this research, we propose paradigms and challenges regarding the motivations and the ethical issues of AI applied to education.

2. Connecting AI with the educational ecosystem: opportunities, challenges, or risks?

Since UNESCO adopted the 2030 Agenda for Sustainable Development, whose key objectives include ensuring equitable access to quality education worldwide (SDG4), technological development and evolution have changed teachers' roles in transmitting knowledge. Technological evolution configures a technological ecosystem, a "metaphor to express a necessary evolution of traditional information systems, with solutions based on the composition of different software components and services that share a set of semantically defined data flows" (García-Holgado & García-Peñalvo, 2019). Part of this evolution is the impact that algorithms and robotization are causing in the field of education, so many "fear that robots driven by artificial intelligence will soon completely replace human teachers" (Spirina, 2018). There is a risk in the making, but also an opportunity. It is more than likely that the e-learning industry (Crisol-Moya et al., 2020) will focus on "swarms of robots as a novel educational tool to focus on precisely those cross-cutting skills that are difficult to explain otherwise" (Global Market Insights, 2021). What is beyond doubt is that the role of the teacher is going through a golden age with AI (Benjamins & Salazar, 2020; Craig, 2018; OECD, 2021; Oliver, 2020; Ramírez-Montoya et al., 2022; Selwyn, 2019; UNESCO, 2021a). In

other fields, it has reduced the demand for workers performing mundane tasks, improved output standards, and enabled humanity to solve complicated health, logistics, and safety tasks that require informed decision-making. In this sense, the trend is that AI will create more jobs than it will eliminate (Fundación Telefónica, 2015). However, many professions, such as teaching, training, and education, undoubtedly require the human mind, not AI. This does not mean that the cognitive capacity of teachers, the continuous flow in the transmission of knowledge among their students, and the tasks of management and content creation can be supported by algorithmic assistants and AI tools.

A teacher's workload often exceeds what is reasonable, since a teacher is expected to monitor students' academic performance, grade homework, prepare lessons and a long list of academic activities, and carry out research activity that requires more dedicated time. Thus, the advance of emerging technologies may be on the way to transforming teaching and learning, leading to a disruption in education, as we know it today. With this in mind, experts agree that AI in education has the mission to help in the planning, personalization, visualization, and facilitation of the learning process.

In this context, recent research suggests that AI will be a great novelty in education and the teaching–learning process, as well as the engine of what is already being called Education 4.0 (Fidalgo-Blanco et al., 2022; Ramírez-Montoya et al., 2022). Some argue that the role of AI will range from initial education (infants) to higher education, including business training. It will enhance the student experience by providing the opportunity to create adaptive learning functions and models with personalized tools (Flogie & Aberšek, 2021). It should be noted that both teachers and students are increasingly immersed in the use of technologies and platforms that optimize, on the one hand, the transmission of knowledge and, on the other hand, the acquisition of new learning. In summary, AI will have different educational roles (Affde, 2021; UNESCO, 2021a). Table 1 describes some of the roles and challenges in which AI is changing the field of education globally.

Table 1. AI roles in online education

Role	Description	Challenges
Universal access	AI can make classrooms available to everyone, including those who speak different languages or those with visual or hearing disabilities. For example, "Presentation Translator", a free add-in for PowerPoint, helps create real-time captions when a teacher is teaching a class.	Teacher training in the use of new educational resources.
Intelligent tutoring systems	AI not only has the power to condense lectures into smart cards and guides but can also be used to tutor students based on their difficulties. For example, air force technicians use an intelligent tutoring system called SHERLOCK to detect electrical system problems in aircraft.	The ethical challenge involved in addressing AI design, with the consideration of verifying the systems' type of response to prevent bias or prejudice.
Virtual facilitators	With trending technology such as gesture recognition, a teacher can be replaced by a robot. AI, 3D gaming, and computer animation are used to create realistic virtual characters and interactions. Even augmented reality is part of this system.	Design of new pedagogical models applicable to the AI of immersive realities.
Intelligent content	Intelligent content can be created from digitized guides, customized learning interfaces, and digital curricula through various media such as audio, video, and online wizard.	Ethical and legal challenges in intellectual property rights.
Collaboration between teachers and AI	The vision for AI in education envisions this technology and teachers working together to get the best outcome for students. For example, collaborating drives the efficiency, personalization, and optimization of administrative tasks.	Development of digital literacy/AI plans.
Content analysis	Educators and content providers gain essential information about learners' progress through intelligent analytics. Through this, the content taught to learners is optimized for a more significant effect.	Ethical challenges in the processing of student data and information.
Out-of-class tutoring support	With AI, tutoring and support programs are becoming more advanced and can effectively help students with homework or test preparation.	Human–machine/computer interaction (HM/CI) paradigm.
Automate administration tasks	Educators spend much time grading homework and exams. AI systems can grade multiple-choice questions and are close to accessing written answers.	Resource optimization.

3. Ethical aspects and principles of AI for the common good

Digitalization opens up new possibilities for education (OECD, 2021). In this line, the evolution of technology and the development of AI present not only opportunities but also risks and, above all, many unknowns and concerns (Selwyn et al., 2022). One of the risks has to do with access to information that is being heavily manipulated by "behavior modification empires" (Lanier, 2018, p. 22) in companies such as Facebook, Google, and Twitter, or how an army of mathematicians is being hired to process information "with data extracted from social networks and e-commerce sites in order to analyze people's behavior" (O'Neil, 2017, p. 10). Furthermore, what proliferates most in education are data and information. With this scenario full of unknowns, challenges to the ethics of AI are configured as the epicenter of the debate.

Therefore, at a more general level, there is the "need to differentiate between doing ethical things and doing things ethically, to understand and to make pedagogical choices that are ethical, and to account for the ever-present possibility of unintended consequences" (Holmes et al., 2022), nuances that are necessary to underline in the implementation of AI in the educational field. In this context, the underlying questions are the following: What is the role of AI in education? What is the truth in the hypothetical replacement of teachers by AI? Can AI cultivate critical thinking in the minds of students? How do (or can) rapid advances in AI, learning analytics, robotics, etc. change the way teachers teach and students learn? (OECD, 2021). Moreover (Selwyn, 2019), should we put the judgments of humans above those of machines? Should systems and applications that prove effective in terms of learning –or saving money– be mandatory? What does the continued rise of AI mean for education? What are the implications for education in the next decade? What problems arise that require rethinking AI and how to implement it effectively in educational contexts? Have the ethical issues involved in ceding the field of education to algorithmic developments been addressed?

The latest studies published, not only in scientific media but also in the general public, seek to make people understand the scope of AI, especially because, as Stephen Hawking stated, "AI can be the best or the worst thing that has happened in the History of Humanity" (Infobae, 2017). Thus, the educational environment is affected by all these AI-generated changes, ranging from preschool stages to higher or graduate levels (Moreno & Pedreño, 2020). This implies, as experts argue, "that the complexity and continuous change of the forces of the new environment requires the rethinking of education from a vision of lifelong learning" (Latorre et al., 2021: 13). However, this rethinking with the digital automation of teaching does not simply consist of the technical aspect of designing, programming, and implementing systems with greater efficiency (Selwyn, 2019).

In this sense, several specialists, proper authorities in this technological field, wonder whether strong AI (SAI) would have actual intelligence, genuine understanding, and creativity, and, consequently, whether it would have an identity, moral integrity, and free will. For Boden (2022), "these are not scientific questions, but philosophical ones [...] some would say it does not matter: what matters is what SAIs will actually do. However, the answers could affect how we relate to them".

For her part, Gabriela Ramos, Assistant Director-General for Social and Human Sciences at UNESCO, with the recent "Recommendation on the Ethics of Artificial Intelligence" (Urbán, 2021), maintains that "what happens is that we always try to persecute [the AI advances], but the whole cycle of AI must be contemplated from research, development, implementation, impact assessment, feedback, development of algorithms, the conceptual framework of algorithms... All these steps require ethical reflection and require controls to be established".

In a similar spirit, Benjamins and Salazar (2020) argue that "since it is not possible to teach ethics to a machine, it will be up to the companies and their developers to demand that the development and use of AI be carried out ethically, under international human rights"', applicable to all fields of knowledge, including education. For its part, the European Commission proposes several guidelines to be considered when developing reliable and safe ethics in the European Union.

These are issues that increasingly affect and concern the academic community: teachers and students, but they also affect organizations, public and private entities, governments, and, in short, society in general. However, as Oliver (2020) points out, the "first step to facing what we fear is to learn; to learn, in this case, with the technology we live with, with the aspiration of understanding the world better, and, therefore, to be able to make informed decisions about it".

In this horizon of learning about technology and the potential of AI (Craig, 2018; Mondada et al., 2017; Nye, 2015), other key technologies that apply to the field of AI should be considered. For example, learning analytics, although still a young field, is a powerful resource for informed decision-making and better learning outcomes. One of the studies in the "Working Papers on Education Policy" (UNESCO, 2019) argues that learning analytics applies different areas of knowledge, such as sociology, psychology, ethics, and pedagogy, and can now access the digital revolution to collect a large amount of data that can be analyzed to extract insights or even develop smart tools useful for educational or administrative tasks. Another example is big data. According to Bonami et al. (2020), "AI has become a natural development of

an intelligent system that needs to deal with Big Data, which is why the terms are structurally connected". Also, one more example is robotics with the design and creation of robot teacher models. Indeed, recent generations of robots in classrooms (Selwyn, 2019) consist of computer systems designed to learn independently, develop logical processes, and formulate mental models that allow them to make complex decisions. Hence, "it is increasingly common for algorithms and programs to perform our functions, to the point that people have become, in a way, dispensable in some jobs" (Cukier et al., 2021), so the educational field does not escape these impacts. In fact, Selwyn (2019) argues that "the long-held professional status of teachers in schools and universities is undoubtedly under threat". To this, we must add the argument made by Cukier et al. (2021), who state that AI algorithms are unable to impose constraints and limitations on their own as people do: "The large computational capacity we have today, machines can evaluate a larger decision space much more efficiently and quickly than a person in the same amount of time".

Therefore, implementing AI technologies in education must be seen as a complex and highly controversial issue (Selwyn, 2019), where morals and ethical principles must be solidly supported through regulation with a global approach.

If the application of AI, in general, raises questions, several researchers (Holmes et al., 2022; Osetskyi et al., 2019; Selwyn et al., 2022) point out that AI, along with its ethical uses in the field of education, generates uncertainty and expectation among the various actors in the academic world, perhaps in the same way as happens with any disruptive technology applied to another field of knowledge. Hence, the conclusions of the UNESCO Forum (UNESCO, 2021a) emphasize that once AI's critical human and technical competencies have been identified, school systems must ensure that all students are best prepared for a world in which AI is ubiquitous.

Experts point out that the many complex challenges associated with AI need to be appropriately considered. For example, it is known that AI is sometimes biased; however, the cause is not always apparent. AI decisions may be biased because the human data on which they are based are biased and/or because the algorithms applied and developed by humans are biased. In other words, AI mirrors and exemplifies human decision-making, which reinforces the need for humans to step up and take responsibility and control. In this regard, an increasing number of researchers (Cukier et al., 2021; Selwyn, 2019) agree with the idea of intelligent systems making decisions and then acting on these decisions, so it obviously has ethical implications. So, the following question arises: what moral and ethical principles should govern what we do with AI in education and what tasks should they be programmed to do on their "own initiative"? The answer (for the time being) is self-evident: the ethical code must be assumed by the creator/programmer of the AI system.

4. Potential and literacy of AI in quality education

The previous step to assuming the principles of ethics in AI requires not only the awareness of values but also more knowledge about the potential of this technology. Therefore, in coherence with the UNESCO Forum, held in December 2020, and SDG4 of the 2030 Agenda, knowledge of AI implies designing and developing an algorithmic literacy plan, which should be included in the educational curricula of any field of knowledge. These interdisciplinary and subject-specific curricula that include learning about AI (from its technological explanation to the ethical and philosophical issues of its impact) should reference what has been done by pioneering countries. Furthermore, they should be flexible, open, inclusive, and continuously evolving. Hence, multiple stakeholders could be involved in the design of learning materials, such as AI-based interactive textbooks. According to the "International Forum on AI and the Futures of Education Developing Competencies for the AI Era" (UNESCO, 2021a), "technology-oriented competencies focus on AI techniques, technologies, and applications, and include the advanced AI knowledge and skills needed to create, manipulate, implement and interpret AI". In this regard, the studies analyzed agree that AI literacy should take both a specific and an interdisciplinary approach: "Specific curricula and courses, covering both human and technological aspects of AI, need to be established, building on existing ICT curricula and courses. In addition, the potential and impact of AI should be considered in all school subjects, be they sciences, humanities, or arts" (UNESCO, 2021a). Nevertheless, the most important thing lies in training the trainer, i.e., teacher training plans,

so that teachers and educators can receive adequate training. In this line, both researchers and experts, as well as institutions and organizations (UNESCO, 2022; IRCAI, 2022), propose an ethics of AI, which should address various issues and fundamental principles based on the responsibility, privacy, fairness, and explainability (Villas & Camacho, 2022) of the same that should be included in a digital literacy plan (García-Orosa, 2021; Salazar & Benjamins, 2021). All agree that it is necessary to provide teachers with skills and digital competencies since today's society, and even the very existence of the human being inserted in it, is increasingly dependent on access to communication and information technologies (Aguaded & Romero-Rodríguez, 2015). The field of education is no stranger to this dependence. AI competencies could also be developed in extracurricular activities, such as seminars, workshops, coding debate clubs, or the realization of hackathons, as well as in lifelong learning programs related to academic, scientific, and professional work (Spirina, 2018). Such integration would equip students to increasingly understand how to interact with AI systems, make informed decisions, and prepare for the societal impact of the widespread use of AI in employment, health, democracy, and, in short, in their daily lives. In this regard, Long and Magerko's (UNESCO, 2022) proposal on "AI Literacy: Competencies and Design Considerations" may be emphasized. They propose a set of competencies and design considerations for AI literacy based on a scoping study of existing research, which sought to determine emerging themes in 1) what AI experts believe a non-technical audience should know and 2) common perceptions and misconceptions among learners.

In this context, current AI-driven educational tools, while effective in some contexts, remain limited. Most are limited by the variety of subjects they cover, generally mathematics and some sciences; by the pedagogical approach they take, often at odds with approaches inspired by learning science orthodoxy; and by their prioritization of machine-based (human–computer) interactions over human-to-human (human-to-human) interactions. Both the potential and limitations have yet to be fully identified or addressed. However, it is clear and widely accepted that AI should not (but perhaps can) replace schools or teachers (Selwyn, 2019). Instead, social interactions, between students and teachers and among students themselves, should remain the epicenter of learning. In addition, some existing AI tools can reduce students' capacity, so it is essential to support this group to be active participants in their learning processes and take more responsibility for their own learning.

5. Results and conclusions

According to our initial approach, from the conclusions drawn, it can be deduced that the scope of SDG4 for Quality and Inclusive Education of the 2030 Agenda, among other goals, entails the access to and continuous use of emerging technologies that generate disruption in the teaching and learning process. AI, as part of that disruption, is increasingly immersed in the field of education, so it can easily become the best ally of students and teachers, providing personalized pedagogical content and personalized tutoring and assistance when needed. Therefore, in addition to their cognitive capabilities in algorithmic design, AI developers must consider ethical principles in developing tools that they design with goals or claims to teach better than teachers. Thus, AI regulations and ethics must be achieved without compromising human values, undermining diversity, and creating new inequalities.

However, it must be made clear that AI cannot be a source of inspiration for learning and that, with a human teacher, where empathy in the teaching-learning process is higher, such inspiration plays a key role. Moreover, this is the main disadvantage of AI-driven education. At the same time, this is the main reason it will not wholly replace human teachers. However, AI is already being used to support remote learning in different modes with remotely delivered, guided, and shared classes, each of which brings benefits and challenges. On the other hand, the various AI models created by Big Tech (GAFAM) (Benjamins & Salazar, 2020) are becoming increasingly sophisticated. However, expert assessments differ regarding whether SAI (for the moment) has real intelligence, understanding, or creativity, but even less identity, moral integrity, or free will. However, advances in AI and machine learning coupled with the development of natural language processing are changing and evolving rapidly, so it is likely that in little more than five years, the "technological singularity" (Benjamins & Salazar, 2020) of AI, i.e., the moment when AI catches up with human intelligence, will occur. All these trends will continue to crescendo, so it is necessary to focus

attention on the ethical implications, especially in a field such as education. Therefore, without claiming to be exhaustive and based on the various reports and references cited, it is suggested to reflect on proposals for some paradigms and ethical challenges of AI in education. These challenges should be posed from both teaching and research perspectives and from the point of view of students and institutions. The resolution of these approaches would help clear the unknowns (raised in Section 3) and blur the impact that AI systems currently present, not only in education but also in society.

5.1. Teaching perspective

The literature review highlights the idea of formulating consensus so that each country's educational systems should define how to use the data of the actors involved in using AI: teachers and students. Likewise, there is a need for an AI literacy plan to train teachers in technical skills and ethical–philosophical debates. This implies that AI will change the role of the teacher, with the expectation that it will take over most of the knowledge-based teaching and assessment, enabling teachers to focus on the social aspects of education. Finally, we should assume that AI could support teaching in various ways: open educational resources, content recommendation (Manrique-Losada et al., 2020), student emotion detection, intelligent tutoring systems, AI-driven teaching assistants, automatic grading of exams, and automatic monitoring of forums. These proposals are imbued with ethical challenges that must be taken up by the various actors in the development of AI, mainly linguists and behavioral scientists.

5.2. Research perspective

First, the objective is to join efforts for the creation and implementation of an ethical observatory of AI in education, with a cross-cutting and multidisciplinary base, whose purpose involves analyzing relevant AI initiatives applied to this field, but, above all, to serve as a platform for the exchange of information, analysis and research on strategic plans, ideas, and guidelines in the proposals for regulations and standardization of an ethical code of AI in/by/for education.

Second, in anticipation of an expanded range of teaching and learning scenarios, it is essential to deepen research in AI and education, developing AI systems that help teachers and improve teaching with responsible, ethical, and equitable AI.

5.3. Students' perspective

Students are at the center of learning, so AI is becoming a catalyst for reforming education under new pedagogical frameworks that will enable a more significant emphasis on project-based learning, flexible learning, collaborative learning, and self-regulated learning, thus improving overall educational quality. Students will have to develop a new range of digital competences around issues such as information processing, computational thinking, and digital learning. For this group, the ethical challenges and vulnerabilities in using AI must be examined, as they underlie some of the risks outlined in Section 3.

5.4. Institutional perspective

Within this context, it is imperative to address and develop consensual public policy frameworks to regulate and raise awareness of the ethical use of AI in education. All this considering that the main obstacle to the widespread adoption of AI technologies is the lack of solid evidence on their effectiveness and impact on students' academic achievement. Hence, it is necessary to improve the governance, accessibility, and reliability of AI, as well as the professional development of teachers.

Authors' Contribution

Idea, J.F., F.G.; Literature review (state of the art), J.F., F.G.; Methodology, J.F.; Data analysis, J.F., F.G.; Results, J.F., F.G.; Discussion and conclusions, J.F., F.G.; Drafting (original draft), J.F.; Final revisions, J.F., F.G.; Project design and sponsorships, J.F., F.G.

Funding Agency

We acknowledge the funding of UCM Groups 2021-22 (Internet MediaLab Ref. 970620), UCM's own Research Support Program of the Vice-Rectorate for Research and Transfer, and AVisSA Project (PID2020-118345RB-I00) by the Ministry of Science and Innovation of the Government of Spain.

References

Affde (Ed.) (2021). *Impacto de la IA en la industria del e-learning para 2021 y más allá*. Affde. https://bit.ly/3Gvqd3f

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. *Education in the Knowledge Society*, *16*, 44-57. https://doi.org/10.14201/eks20151614457

Alonso-De-Castro, M.G., & García-Peñalvo, F.J. (2022). Successful educational methodologies: Erasmus+ projects related to e-learning or ICT. *Campus Virtuales*, *11*, 95-114. https://doi.org/10.54988/cv.2022.1.1022

Benjamins, R., & Salazar, I. (2020). *El mito del algoritmo. Cuentos y cuentas de la Inteligencia artificial*. Ediciones Anaya Multimedia. https://bit.ly/3QfzhwR

Boden, M.A. (2022). *Inteligencia Artificial*. Turner Publicaciones S.L. https://bit.ly/3zVqBqo

Bonami, B., Piazentini, L., & Dala-Possa, A. (2020). Education, Big Data and Artificial Intelligence: Mixed methods in digital platforms. [Educación, Big Data e Inteligencia Artificial: Metodologías mixtas en plataformas digitales]. *Comunicar*, *65*, 43-52. https://doi.org/10.3916/C65-2020-04

Canavilhas, J. (2022). Inteligencia artificial aplicada al periodismo: estudio de caso del proyecto «A European Perspective» (UER). *Revista Latina de Comunicación Social*, *80*, 1-13.

Craig, S. (2018). Tutoring and intelligent tutoring systems. *Nova Science Publishers*.

Crisol-Moya, E., Herrera-Nieves, L., & Montes-Soldado, R. (2020). https://doi.org/10.14201/eks.20327

Cukier, K., Mayer-Schönberger, V., & De-Vericourt, F. (2021). *Framers. La virtud humana en la era digital*. Turner Publicaciones S.L. https://bit.ly/3OWbocD

C. Europea (Ed.). (2020). *Libro Blanco sobre la inteligencia artificial: Un enfoque europeo orientado a la excelencia y la confianza*. https://bit.ly/3aTX8n2

Fidalgo-Blanco, A., Sein-Echaluce, M.L., & García-Peñalvo, F.J. (2022). Método basado en Educación 4.0 para mejorar el aprendizaje: Lecciones Aprendidas de la COVID-19. *RIED*, *25*(2), 49-72. https://doi.org/10.5944/ried.25.2.32320

Flogie, A., & Aberšek, B. (2021). *Artificial intelligence in education*. IntechOpen. https://doi.org/10.5772/intechopen.96498

Flores-Vivar, J.M., & Zaharía, A.M. (2019). Las redes sociales académicas en la difusión de la producción científica. *Estudios sobre el Mensaje Periodístico*, (3), 25-25. https://doi.org/10.5209/esmp.66993

Fundación Telefónica (Ed.) (2015). *El trabajo en un mundo de sistemas inteligentes*. Ariel. https://bit.ly/3PWQlrQ

García-Holgado, A., & García-Peñalvo, F.J. (2019). Validation of the learning ecosystem metamodel using transformation rules. *Future Generation Computer Systems*, *91*, 300-310. https://doi.org/10.1016/j.future.2018.09.011

Global Market Insights (Ed.) (2022). *Industry trends*. Elearning Market Report. https://bit.ly/3LZbT4l

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S.B., Santos, O.C., Rodrigo, M.T., Cukurova, M., Bittencourt, I.I., & Koedinger, K.R. (2022). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education*. https://doi.org/10.1007/s40593-021-00239-1

Infobae (Ed.) (2017). *a cruda advertencia de Stephen Hawking sobre el desarrollo de la inteligencia artificial*. Infobae. https://bit.ly/38XreF4

IRCAI (Ed.) (2022). *IRCAI Global Top 100 Report 2021*. UNESCO. https://bit.ly/3J1qb4k

Kiecza, D. (2022). Classroom: Practice sets: A more personal path to learning. *The Keyword Google*. https://doi.org/10.52188/jpfs.v5i1.224

Lanier, J. (2018). *Diez razones para borrar tus redes sociales de inmediato*. Penguin Random House Grupo editorial. https://bit.ly/3Qljpct

Latorre, C., Sierra, V., & Lozano, R. (2021). *El docente del siglo XXI. Enfoques y metodologías para la transformación educativa*. Prensas de la Universidad de Zaragoza. https://bit.ly/3aQZEdD

Manrique-Losada, B., Zapata-Cárdenas, M.I., & Arango-Vásquez, S.I. (2020). Entorno virtual para cocrear recursos educativos digitales en la educación superior. *Campus Virtuales*, *9*, 101-112. https://bit.ly/3zAmM8G

Mondada, F., Retornaz, P., & Bonani, M. (2017). Bringing robotics to formal education: The Thymio open-source hardware robot. *IEEE Robotics & Automation Magazine*, *24*(1), 77-85. https://doi.org/10.1109/MRA.2016.2636372

Moreno, L., & Pedreño, A. (2020). *uropa frente a EE.UU. y China. Prevenir el declive en la era de la inteligencia artificial*. KDP. https://bit.ly/3PFeOS2

neil, C.O. (2017). *Como el Big data aumenta la desigualdad y amenaza la democracia*. Capitán Swing Libros. https://bit.ly/2Y0njfS

Nye, B.D. (2015). Intelligent tutoring systems by and for the developing world: A review of trends and approaches for educational technology in a global context. *International Journal of Artificial Intelligence in Education*, *25*(2), 177-203. https://doi.org/10.1007/s40593-014-0028-6

OECD (Ed.) (2021). *OECD Digital Education Outlook 2021: Pushing the frontiers with artificial intelligence, blockchain and robots*. OECD. https://doi.org/10.1787/589b283f-en

Oliver, N. (2020). *Inteligencia Artificial, naturalmente. Centro de Publicaciones*. Ministerio de Asuntos Económicos y Transformación Digital. https://bit.ly/3Olt5SE

Osetskyi, V., Vitrenko, A., Tatomyr, I., Bilan, S., & Hirnyk, Y. (2019). Artificial intelligence application in education: Financial implications and prospects. *Financial and Credit Activity Problems of Theory and Practice*, *2*, 574-584. https://doi.org/10.18371/fcaptp.v2i33.207246

Parlamento Europeo (Ed.) (2021). *La inteligencia artificial en los sectores educativo, cultural y audiovisual*. https://bit.ly/3oldM1M

Ramírez-Montoya, M.S., Castillo-Martínez, I.M., Sanabria-Z, J., & Miranda, J. (2022). Complex thinking in the framework of education 4.0 and open innovation-a systematic literature review. *Journal of Open Innovation*, *8*(1).

https://doi.org/10.3390/joitmc8010004
Salazar, I., & Benjamins, R. (2021). *El algoritmo y yo. GuIA de convivencia entre seres humanos y artificiales*. Ediciones Anaya Multimedia. https://bit.ly/3BU2VUL
Selwyn, N. (2019). *¿Debería los robots sustituir al profesorado? La IA y el futuro de la educación*. Ediciones Morata. https://bit.ly/3zxyPmO
Selwyn, N., Rivera-Vargas, P., Passeron, E., & Miño-Puigcercos, R. (2022). ¿Por qué no todo es (ni debe ser) digital? Interrogantes para pensar sobre digitalización, datificación e inteligencia artificial en educación. In P. Rivera-Vargas, R. Miño-Puigcercos, & E. Passeron (Eds.), *Educar con sentido transformador en la universidad* (pp. 137-147). Octaedro. https://doi.org/10.31235/osf.io/vx4zr
Spirina, K. (2018). *Is AI here to replace human teachers or is it a teacher's assistant?* Towards Data Science. https://bit.ly/3N27Xkq
Túñez-López, J.M., & Tejedor-Calvo, S. (2019). Inteligencia artificial y periodismo. *Doxa Comunicación*, *29*, 163-168. https://doi.org/10.31921/doxacom.n29a8
UNESCO (Ed.) (2019). *Artificial Intelligence in education: Challenges and opportunities for sustainable development*. Unesco Working Papers on Education Policy. https://bit.ly/3z6BQvN
UNESCO (Ed.) (2021a). *International Forum on AI and the futures of education developing competencies for the AI era*. UNESCO. https://bit.ly/3zoB6AS
UNESCO (Ed.) (2021b). *Recomendación sobre la ética de la Inteligencia Artificial*. UNESCO. https://bit.ly/3lTIvSf
UNESCO (Ed.) (2022). *K-12 AI curricula: A mapping of government-endorsed AI curricula*. UNESCO. https://bit.ly/3B6f6xi
United Nations (Ed.) (2019). *The Sustainable Development Goals Report 2019*. United Nations. https://bit.ly/34nbq60
Urbán, A. (2021). *UNESCO busca regular éticamente a la Inteligencia artificial*. El Universal. https://bit.ly/3PQ4jMl
Villas, M., & Camacho, J. (2022). *Manual de ética aplicada en Inteligencia Artificial*. Anaya. https://bit.ly/3vx7kZF
Vitanza, A., Rossetti, P., & Mondada, F. (2019). Robot swarms as an educational tool: The Thymio's way. *International Journal of Advanced Robotic Systems*, *16*(1). https://doi.org/10.1177/1729881418825186

Comunicar, n. 74, v. XXXI, 2023 | Media Education Research Journal | ISSN: 1134-3478; e-ISSN: 1988-3478
www.comunicarjournal.com

Big Data and Business Intelligence on Twitter and Instagram for digital inclusion

Big Data y Business Intelligence en Twitter e Instagram para la inclusión digital

Carlos Barroso-Moreno. Predoctoral Research Fellow, Department of Educational Studies, Complutense University of Madrid (Spain) (carbarro@ucm.es) (https://orcid.org/0000-0002-1609-2267)
Dr. Laura Rayon-Rumayor. Professor, Department of Educational Studies, Complutense University of Madrid (Spain) (larayon@ucm.es) (https://orcid.org/0000-0003-0339-8221)
Dr. Antonio Bautista García-Vera. Full Professor, Department of Educational Studies, Complutense University of Madrid (Spain) (bautista@edu.ucm.es) (https://orcid.org/0000-0002-5194-7419)

ABSTRACT

Social media can contribute to an inclusive society, but they are also asymmetrical and polarised communication spaces. This requires competent teachers to build critical digital citizenship. The aim of this article is twofold: to present web scraping and text analytics as tools that define teachers' digital competences, and to investigate which posts on Twitter and Instagram are most viral in relation to education, disability and inclusion. A total of 48,991 publications in Spanish and English were analysed, corresponding to the period from 13 October 2021 to 1 May 2022. The 100 most viral posts were selected, and correlations were identified between the sentiment, gender and influence associated with the content, its temporal and geographic space. The results show that economic and political influence groups are the most viral, relegating non-profit organisations or individuals with altruistic outreach to second place; only on international days is this trend reversed. Bots do not interfere to impose messages; it is artificial intelligence algorithms that overshadow vindictive and humanistic content. The most influential people are predominantly male, associated with institutional accounts in the political sphere. It is concluded that Big Data and Business Intelligence tools help teachers to analyse relevant educational and social issues, and to acquire a collective ethic in the face of new educational challenges.

RESUMEN

Las redes sociales pueden contribuir a una sociedad inclusiva, pero también son espacios de comunicación asimétricos y polarizados. Ello requiere de un profesorado competente para la construcción de una ciudadanía digital crítica. Este artículo tiene un doble objetivo, presentar las técnicas «Web scraping» y «text analytics» como herramientas que definen competencias digitales docentes, e indagar sobre qué publicaciones, en Twitter e Instagram, son más virales en relación con educación, discapacidad e inclusión. Se analizaron 48.991 publicaciones en español e inglés, correspondientes al periodo del 13 de octubre de 2021 al 1 de mayo de 2022. Se seleccionaron las 100 más virales, e identificaron las correlaciones entre el sentimiento, género e influencia asociado al contenido, su espacio temporal y geográfico. Los resultados evidencian que los grupos de influencia económica y política son los más virales, relegando a un segundo plano a las organizaciones sin ánimo de lucro o particulares con difusión altruista; solo en los días internacionales se invierte esta tendencia. Los «bots» no interfieren para imponer mensajes, son los algoritmos de inteligencia artificial los que opacan contenido reivindicativo y humanístico. Las personas más influyentes tienen una prevalencia de género masculino asociadas a cuentas institucionales de ámbito político. Se concluye que las herramientas de «Big Data» y «Business Intelligence» ayudan al profesorado a analizar temas educativos y sociales relevantes, y a adquirir una ética colectiva frente a los nuevos retos educativos.

KEYWORDS | PALABRAS CLAVE

Social network analysis, Big Data, education, disability, digital inclusion, influence groups.
Análisis de redes sociales, Big Data, educación, discapacidad, inclusión digital, grupos de influencia.

Received: 2022-05-29 | Reviewed: 2022-06-21 | Accepted: 2022-07-13 | OnlineFirst: 2022-10-01 | Published: 2023-01-01
DOI https://doi.org/10.3916/C74-2023-04 | Pages: 45-56

1. Introduction and state of the art

Current reports (We are social, 2021; Ditrendia, 2020) point to the exponential increase of users connected to social networks worldwide, being not only individuals, but also professional and institutional groups and media that impact on the construction of reality and how meanings are shared (Dellwing, 2021; Del-Fresno-García, 2014, 2019; Ladogina et al., 2020).

In terms of studies on the functions, roles and relationships between members (Awidi et al., 2019; Brunner et al., 2019; De-Groot et al., 2022; Grace et al., 2019; Tuzel & Hobbs, 2017; White & Forrester-Jones, 2020), we are particularly concerned with the approach that emphasises social networks as asymmetric spaces of communication. Specifically, Barberá (2015), Barberá et al. (2015) and Brady et al. (2019) point out that users share messages that represent beliefs, opinions and values that they endorse and follow profiles they trust ideologically. Another relevant factor that conditions communication is the moral-emotional language that political leaders use and its moral contagion effect, as well as the asymmetry in communication between content creators and their potential followers, depending on the content of the messages, an asymmetry amplified by bots (Robles et al., 2022).

There is no doubt about Twitter's potential for certain users and groups, such as political elites and corporations, to reach large audiences, potential voters and consumers, through direct and indirect links of influence that define interaction on this network, a predictive factor of influence and social impact outside the network (Brady et al., 2019). In this context, hate speech is of particular relevance, as it is generating significant polarisation among people based on their ideology, with effects on the "selective perception bias" that favours the positive evaluation of the message of issuers with whom there is ideological affinity and the rejection of speeches with an opposing ideology. This is an important predictive factor for behaviour towards certain offline groups, which are represented through exclusionary and anti-democratic messages on social media (Ortega-Sánchez et al., 2021).

To address this concern, we understand that teachers have a key role to play in building a digital citizenship with the capacity to participate in an informed and responsible way, and thus contribute to a democratic and inclusive networked society (Bautista, 2021; Carlsson, 2019; Ortega-Sánchez et al., 2021). As evidenced by Tuzel and Hoobs (2017: 64), the use of social media for intercultural citizenship requires teachers to have a "solid appreciation of the asymmetries and inequalities inherent in information flows", and an understanding of how these digital platforms function as spaces for dissemination, amplification of ideas and mobilisation of actions for groups and individuals with unequal rights, such as those with disabilities, in order to disseminate and critique ideas, as well as publicise their achievements for a more inclusive society (Hemsley et al., 2018).

So, how should teachers work on these digital competences that prepare them for inclusive education? What tools and procedures will help them in the complex and beautiful task of knowing the meanings, beliefs and attitudes that circulate on social networks with thousands and thousands of participants? We understand that an appropriate way to investigate and generate the knowledge that teachers must have in order to promote relationships that lead to creating feelings of inclusion and belonging of students to the reference group, is through the tracking of social networks with techniques and tools such as Web Scrapping and text analytics that we present in this article.

In view of the above, concerns and proposals derived from the review of the state of the art, this article has a dual purpose. One, to exemplify in a research context the use of these tools that make up one of the digital teaching competences to encourage debate, and to inspire and illuminate evidence of the value of the techniques mentioned in the analysis of social media and interactive data visualisation. The other is to answer three questions/hypotheses on the processes of asymmetric relationships that help to better understand the processes of remodelling and relegation of these subjects on social media, useful knowledge for defining the content of the digital teaching competence in particular:

- H1: The influence of economic and political power groups together with an interest of digital platforms reshape the issues associated with education, disability and inclusion according to their own interests.
- H2: Artificial intelligence algorithms relegate minority ideas or altruistic broadcasts to second or third place.

- H3: Posts with negative sentiment are associated with political groups and leaders, compared to positive posts linked to associations and individual content creators advocating for the human rights of persons with disabilities, with no influence by gender.

An innovative contribution of the article is the access to the data and results obtained, which are made available to researchers, teachers and other professionals so that they can dynamically and interactively read them.

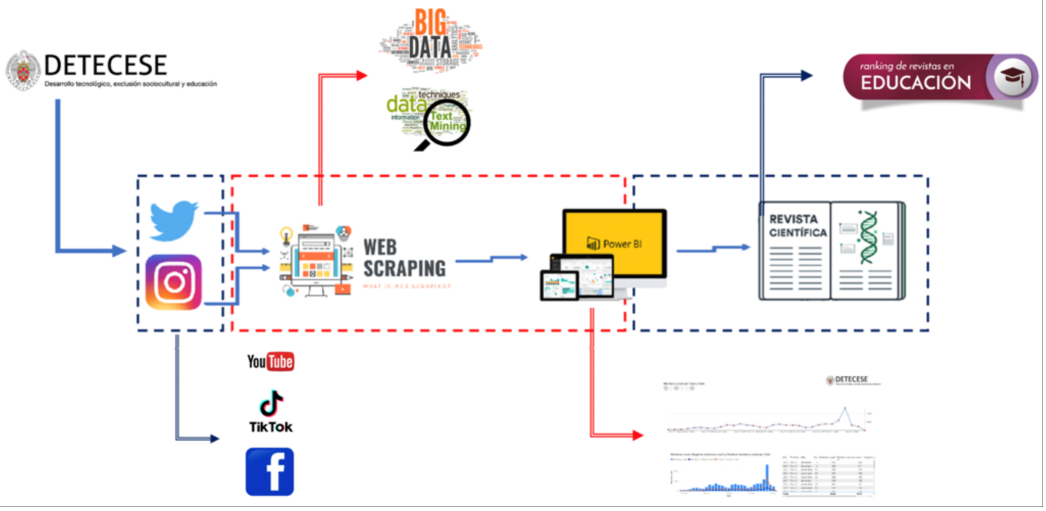
2. Material and methods

In this article we analyse which posts on Twitter and Instagram are most viral in relation to education and disability and inclusion: in this sense, it will allow us to better identify correlations between content, sentiment, gender, influence, and their temporal and geographic space. The database is composed of posts downloaded from both social networks on education content related to inclusion or disability over 200 days. Social Big Data Analysis techniques are applied, such as Web Scraping techniques to extract the information and analyse it with Big Data and Text Mining algorithms. Subsequently, the results are represented using the Power BI business tool, allowing readers to interact dynamically.

2.1. Process flow in the methodology

The set of research processes, from the origin of the data to the graphic representation, is conceptually referred to as the "Social Networks Tools", which is represented in Figure 1. The processes are divided into two large blocks; those delimited in blue are modifiable by the researcher, such as the origin of the data or the keywords used. The block delimited with red dashed lines refers to Web Scraping techniques, Big Data algorithms, Data Mining and part of the Business Intelligence tool, belonging to the research group with intellectual property, in which the interested parties cannot make modifications or have full access.

Figure 1. Flow of the methodology used in the Social Networks Tools tool



The first process corresponds to the selection of social networks with the possibility of data extraction, such as Twitter, Instagram, YouTube, TikTok or Facebook. At this point, the necessary keywords are selected in order to consider the storage of the information. The second process applies Web Scraping techniques for data extraction in social media with Python, which literally means scraping from the web (Mitchell, 2018). Subsequently, Big Data techniques are applied with the information obtained to process the large amount of data collected, and data mining is used to analyse text from the publications. Then, variables specific to the research on the communication profile such as gender, professional activity, geolocation and temporal space are included. The next process is to load the data into the Microsoft Power BI data analysis service in order to provide interactive visualisations for researchers and faculty in order to generate their own reports and dashboards (Becker & Gould., 2019). Finally, the user interprets

the generated graphs for the understanding of the subject matter addressed, in search of a scientific and/or educational use, with a real impact on society.

Extract, Transform and Load (ETL) processes allow data from multiple sources to be mixed. In our particular case, data from Twitter and Instagram is extracted, transformed by Python script with a series of requirements and loaded into a data warehouse to channel the data in a homogeneous way and analyse it using different algorithms.

2.2. Data extraction

Twitter data is downloaded from the Twitter API using the "tweepy" library and Instagram data is downloaded using private algorithms developed specifically for the research. In both cases, real-time data is collected using the Python programming language. Given the need to capture all traffic in real time, unlike most previous studies, we use posts that have the demanded keywords; we do not focus on what a particular group of profiles are saying to avoid bias. This has another implication: the data capture starts at the moment of the activation of these words; we do not collect data retroactively.

The database generated consists of publications in the period from 13 October 2021 to 1 May 2022, with a total of 48,991 publications in the Spanish and English-speaking world. The requirement is to contain a number of keywords regardless of the user disseminating them. The common keyword used is education, and at least one disability or inclusion word, as well as words derived from them. In addition, words without accents are included due to the virality of some news items with such spelling. Therefore, any publication that has the word 'education' together with one or more words for disability or inclusion on the social network Twitter or Instagram is considered. The choice of the word 'education' together with 'disability' or 'inclusion' is motivated by the fact that certain publications that deal with disability do not refer to this term, but to the idea of inclusion; in addition, inclusion does reflect different social sensitivities around the topic addressed. The publications are updated with all the necessary information one week after publication; this situation does not significantly affect the data because the publications have a significant impact in the first days of dissemination, which is empirically proven. In fact, the reader can corroborate in later sections how the number of likes, followers, followings or comments does not increase significantly; therefore, this situation does not influence the subsequent analysis.

2.3. Data transformation

Once the publications were stored in raw form, data mining was performed with Natural Language Processing algorithms from the NLTK and Scikit-Learn Python libraries, a branch of artificial intelligence to determine the interaction between computers and humans (Cheng & Tsai., 2019). At this point, two transformations are differentiated: manual and automatic. The automatic transformations are applied to the 48,991 research publications; the manual transformations are applied by the research team to the 100 publications of greatest interest with particular classifications for the subject matter addressed. All the variables available in the database are detailed by group, type, category and example (Table 1).

2.3.1. Automatic

The application of text mining begins with the process of tokenisation of the content, which allows words to be separated by the spaces that make up the sentence. Next, the words known as "stopwords" are eliminated, consisting of prepositions, determiners or particular words, among others. This separation and cleaning of the text allows for the analysis of repetition frequency, word clouds, etc. Regarding sentiment analysis, Liu's (2010) dictionary is applied to detect the positive, negative or neutral content of words, providing a final value to the sentence as a whole. Among the limitations of this type of dictionary is the invisibility or misclassification of ironies or puns. In reference to the detection of thematic topics, tweets are grouped by the hashtags for Instagram content, depending on the most frequent words, so that they can be grouped under one of the hashtags. Based on bot detection, the Botometer API for Twitter profiles is used to extract more than 1,200 features such as activity patterns, language, sentiment, social structure or friends, assigning a 1 if it is a bot or a 0 if the account is real. High criteria are set to consider it a bot and not to include real users in that category. Therefore, a cross validation for the Area Under the ROC Curve (AUC) of 0.99 is established.

2.3.2. Manual

In order to answer the hypotheses, put forward, it is necessary to generate specific variables, which cannot be automated. The research team distributes the labelling with homogeneous criteria to ensure a correct classification. This requires a thorough analysis of user profiles, based on tracking publicly disseminated information on other networks or personal blogs to determine variables such as: gender, field, profession, multimodality or estimated locations.

2.4. Data upload

To analyse the data, once the data mark, known as education, is available, with all the variables cleaned, the data is loaded into the Power BI programme. The objective is to analyse the data and indicators to test hypotheses through different visual data analysis. Some of the graphs deal with simple and complex statistics, including maps (geographical or heat maps) on the locations of the most viral senders (Arcila-Calderón et al., 2022), statistical correlation to analyse whether there is a strong or weak relationship, multidimensional graphs, time series or word clouds, among others.

Table 1. Description of the variables, type and categories of the SSR education Data Mart

Group of variables	Variables	Categories	Type	Examples
Identifier	Identifier	-	Numeric	02186348225, 31186048213...
	Label		Categorical	T associated with Top 10 V associated with viral 10 to 50 P associated with publications 50 to 100
	Profile name	-	Text	Mzavalagc
	Real name	-	Text	Margarita Zavala
Transactional data	Origin	-	URL	www.instagram.com/p/CZKkN4lp13G/
	Domain		Categorical	Twitter or Instagram
Temporary	Day	200	Date	2022-01-24
	Time	1440	Date	21:32
Content	Title	-	Text	Real inclusion
	Content	-	Text	Inclusion is achieved through...
	Hashtags	-	Text	#education #education #inclusion
Virality	Retweets (twiitter) Comment (Instagram)	-	Numeric	1957, 34...
	Likes	-	Numeric	33971, 4664...
	Reproductions (Video)	-	Numeric	139103, 83400...
	Following	-	Numeric	826, 453...
	Follower	-	Numeric	1596670, 320...
	Publications (Instagram)	-	Numeric	6068, 234...
Behaviour	Sentiment		Categorical	Positive, neutral and negative
	Profession	>10	Categorical	Politician, digital creator, teacher...
Scope	Scope		Categorical	1.Institutional 2. Group 3. Individual
	Subarea		Categorical	11 Local, 12 Supranational (political), 13 Media 21 Academies, clinics, or companies 22 Non-profit associations 23 Science communicators 31 Professional 32 Professional with disabilities 33 Particular
Multimodality	Modal		Categorical	Single-modal and multimodal
	Submultimodal		Categorical	Text and video; text and image; text audio and audio and image
Demographic	Gender		Categorical	Male, female and neutral
	Estimated location of the person	195	Category	United States, Spain

3. Analysis and results

The publication of the data can be found in the Power Bi tool, which should preferably be opened with the Microsoft Edge browser from a computer (https://bit.ly/3z9wDU6), or with the QR Code, presented in Figure 2.

We extract an initial snapshot of the publications (N=48,991), which are distributed on Twitter with 59.38% (N=29,095) and Instagram with 40.61% (N=19,896). Twitter is the most popular, but Instagram is the most popular in terms of the number of likes and comments.

The results of the analysis of the most viral publications, sentiment and the type of associated profile, show five main ideas presented in Figure 3. (i) The publications with polarised sentiment are the most viral, specifically the positive ones, and of the 10 most viral, 9 are positive. (ii) The 25 most viral publications are mainly campaigns by relevant power groups, orchestrated by multinationals, politicians,

influencers or digital creators with economic interests. Proof of this is the most viral communications [T1] (https://bit.ly/3z5bEBO), [T4] and [V11] related to McDonald's, located in accounts that do not claim any social or educational aspect in relation to disability and inclusion, published on the same date, with the same image and text content.

Figure 2. QR code for access to Power BI

DETECESE

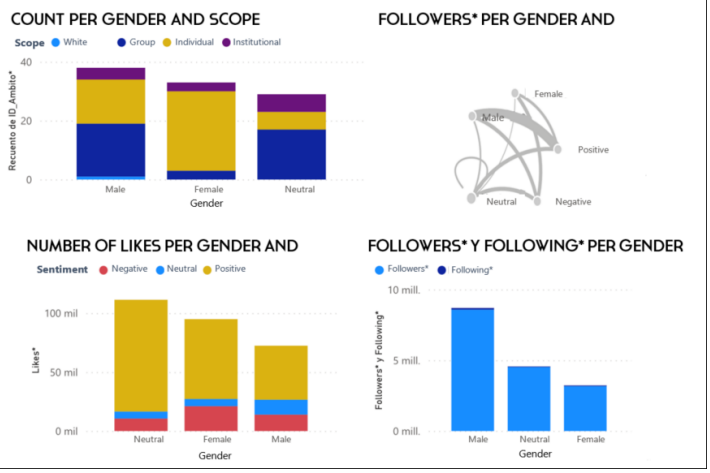
To orchestrate such an impactful campaign requires accounts with thousands of followers; even if it has a low diffusion rate per follower the total impact is very high. (iii) We have to relegate ourselves to position 6 [T6] and 28 [V28] by number of likes to find a social media space. However, [T6] is defined as a space for learning about diversity and LGBTQ+; however, some exclusive content requires a financial outlay (https://bit.ly/3N0Vv4G). One has to go down to position 34 [T34] to find the Only You Are Missing Foundation, a non-profit Non-Governmental Organisation (NGO) for autism spectrum awareness (https://bit.ly/3t2Gpno). (iv) Cases of denunciation of situations of exclusion are almost non-existent and do not have a presence on social networks, mainly due to denunciation from networks with a small number of followers, and a neutral and institutional feel to the denunciations. (v) Mostly, the most disseminated messages are positive or negative, but neutral messages are not so widespread. In light of these results, we can affirm that the hypothesis (H1) is supported: certain groups and political leaders reshape the issues analysed for their own interests, relegating groups and individual actors with altruistic interests to the background.

Figure 3. Number of likes and sentiment on the top 25 most viral posts

The results of analysing the 100 most viral posts from a gender perspective are shown in Figure 4, and are distributed as follows: 38% male, 33% female and 29% neutral. Significant results are as follows: (i) Male accounts have more followers and their comments are harsher (negative) than those of the female gender, the latter being associated with positive comments. In figures, the male gender has 8,582,921 followers and 118,981 following; the female gender has 4,523,345 followers and 43,307 following. (ii) Female accounts have fewer followers, but more likes. (iii) Institutional accounts disseminate more neutral or positive messages. (iv) The male gender has the most institutional accounts on social media. Therefore, hypothesis (H3) is only partially confirmed in that negative sentiments are linked to influential groups

and political leaders, but these accounts are male, which is evidence of a gendered influence. Figure 4 shows the correlation between the variables gender, domain, feelings and multimodality. The results are as follows: (i) Through the nodes we observe a relationship (thick thickness) between the institutional groups and the male gender. The female gender has less presence in the institutional sphere, it appears linked to individual profiles (thin thickness). (ii) Nicolás Maduro has 4 million followers, a paradigmatic case that contributes to the association between the male gender and institutional variables, as happens with other political leaders. The publication associated with this case [P80] compiles the publication (https://bit.ly/3N0QOro), and stands out for its low dissemination ratio with 824 likes in relation to its high number of followers. (iii) Most are multimodal messages (text and image) and monomodal publications are scarce.

Figure 4. Statistics of the 100 most viral posts from a gender perspective



With regard to the detection of bots, with a probability of over 98% that they correspond to automated accounts, 23 accounts were detected. This represents less than 0.07% of the total number of profiles corresponding to 89 posts that were deleted. These publications have themes focused on political issues with a load of negative sentiment in favour of or against certain ideologies and political parties during the election campaign period. Another worrying aspect, as is the case with all social media, is fake news. This is the case of the publication [P60], associated with the Instagram account "adhd_understood" with a significant number of followers, whose owner is Ms. Donna Giachino, who defines herself as a doctor specialising in ADHD (https://bit.ly/3z5ylFV). However, the College of Physicians in Vancouver in 2019 had to rule that she was not registered and therefore could not practice as a speech and language professional (https://bit.ly/3wTC7Rt).

Figure 5 exemplifies the results presented so far, by means of two publications in the context of the last electoral campaign in Chile, published in a close temporal space. They show how the male gender compared to the female gender has a quantitative difference of 560,321 followers and 80,002 following, in line with the gender results (i) of a greater number of followers and following in the male gender. Another relevant aspect of the publications analysed is the number of likes: a publication by José Antonio Kast [V32] has 3,019 likes and 1,265 retweets compared to the 3,854 likes and 3,681 retweets of Claudia Aldana [V17]. In other words, with a lower number of female followers, the publication achieves a greater social impact on Twitter. At this point, the secondary idea of gender (ii) is exemplified as it is women who have more loyal followers in the interaction with the publications. If we look at the number of views of the video [V17], it stands at 85,693 views, a figure higher than the number of followers. These results show that it is the algorithms of the digital platforms themselves that suggest publications to other users, even if they are not followers of the profile, motivated by the polarisation of sentiment associated with electoral campaigns, which leads to segmented loyalty. This situation evidences that algorithms generate echo chambers. In other words, users are presented with suggestions based on their thematic interests and

ideological affinity on a recurring basis, and this procedure is repeated by listening only to information that the digital platforms consider to be of interest to the user. This manipulation of information dissemination leads us to the cognitive biases of the users analysed in the Power BI tool.

Figure 5. Twitter posts related to the Chile 2021 electoral election

Digital platforms play a relevant role in the educational and social sphere because they are a loudspeaker for international days whose purpose is to raise awareness of social issues. As can be seen in Figure 6, the count of publications per day according to sentiment shows the following results. A series of days with the highest number of publications can be seen, which are reflected in the graph in the peaks listed. If we look at the largest, it corresponds to 3 December 2021, the International Day of Persons with Disabilities, with 1,329 publications of which 702 and 199 are of positive and negative sentiments respectively. If we analyse the publications from that day, the most viral publication is positive (https://bit.ly/3GwET20) [T7], associated with a profile outside the traditional political and economic interest groups mentioned above. The creator of this publication is a person with Down's Syndrome, María Jose Paiz Arias, known on social media as Majo. Her publication has 7,034 likes and 55,463 reproductions, whose title conceptualises the vindictive idea of the international days analysed: "inclusion is achieved with fewer labels and more action". However, political leaders such as the Senator of the Republic of Mexico, President of the Commission on the Rights of Children and Adolescents with more than 100,000 followers, only received 13 likes and 6 retweets (https://bit.ly/3wYaCVJ).

In other words, the viral power of influential groups over minority groups is reversed, unlike on other days of the year. Other notable days that show the reversal of the trend are 4^{th} of January World Braille Day, 13^{th} of January World Day to Combat Depression, 26^{th} of January World Environment Education Day and 21^{st} of March World Down Syndrome Day. In reference to World Education Day, 24^{th} of January 2022, the situation described above occurs: political leaders publish, but are relegated to the background regardless of their thousands or millions of followers, as is the case with the official account of the Government of Spain with 778,858 followers, which barely receives 74 likes (https://bit.ly/3lUDyIV). Another significant day, of relevance for this research because it does not directly contain the keywords determined for the study, is the 2nd of April, World Autism Awareness Day, where the most viral publication has a clearly altruistic component from an individual profile, a parent of a child with Autism Spectrum Disorder (ASD): https://bit.ly/3tqQY3R [P95]. In reference to the frequency of publications, weekdays stand out as the time periods of greatest intensity, but on Saturdays and Sundays for the topic analysed, publications are drastically reduced; the causes are not known at the moment. However, it is worth mentioning that, if the international day is a public holiday, it remains positioned as a loudspeaker for demands, due to the high number of publications.

Therefore, it is evident that: (i) international days give visibility to social demands, invisible during the rest of the year in the eyes of social media users, in order to raise awareness, guide and vindicate the social cause. (ii) The profiles involved in social demands are the most viral, relegating influence groups to second

place, although these groups commemorate these days with publications. Therefore, the hypothesis (H1) posed of the dominant prevalence of influential groups has an exception on international days.

Figure 6. Timeline of post counts per day with positive, neutral and negative sentiment, associated with geolocation heat maps of the 200 days of collection

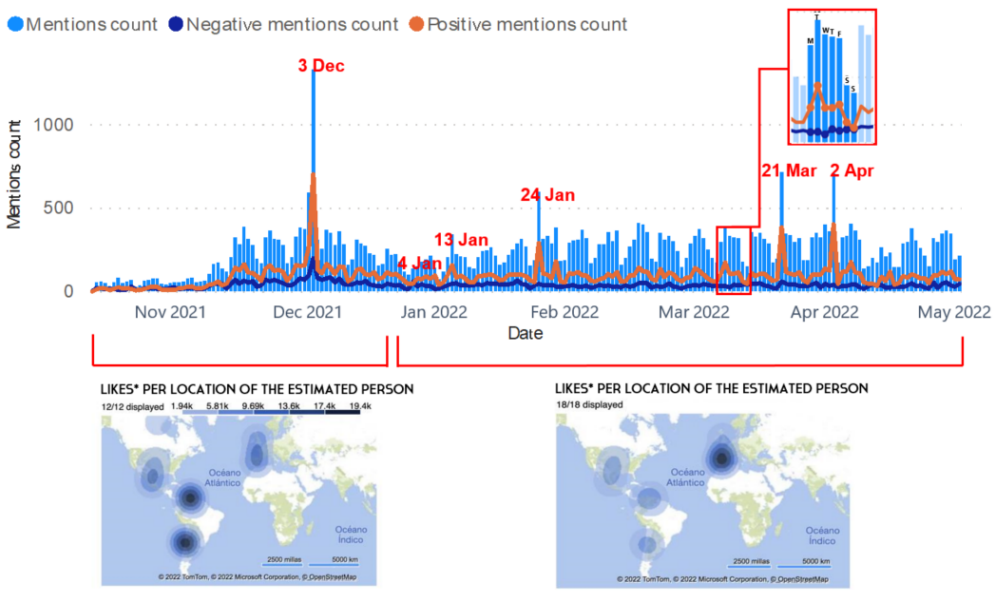


Figure 6 shows a chronogram, associated with heat maps by location of the most viral profiles, which shows the following results: the chronological division to generate these maps is motivated to argue hypothesis H1, where the interests of the influential groups are superimposed on the altruistic groups. If we look at the heat maps generated by the geolocations of the publications, in the dates corresponding to 2021, Argentina and Chile stand out. The reason for this is because the 12^{th} of November was the first round of elections in Chile and on 14^{th} of November the elections in Argentina. In 2022, the focus is on Europe, where different electoral events take place in Portugal, France and Germany. Therefore, (i) events of political interest also overshadow social demands and altruistic content.

If we look at the 100 most mentioned Instagram hashtags (#) in the whole time frame, represented in Figure 6, the first positions are obviously established by the key words of the research. The following hashtags, in the eighth position, are the word autism, #autism (N=2,117) and #autism (N=1,962). According to the World Health Organisation (WHO), the average prevalence of autism is estimated at 1/160 people, therefore, the presence of the topic is logical. Other hashtags reflect #adhd (21) N=913, #dyslexia (23) N=876 #tea (29) N=801 #downsyndorme (35) N=733 or #asperger (86) N=457. In addition, hashtags related to culture, values and respect such as #diversity (10) N=1927, #culture (19) N=1024, #teacher (31) N=773, #specialeducations N=762 stand out. Other hashtags reflecting the tuning of the most social posts in the top 100 are #diversity #equality #equality #values #equity #respect #inclusiveeducation #accessibility. (i) Aligned with the H1 of influencers among the most repeated hashtags is #influencer, related to digital creators. (ii) In line with the philosophy that education professionals play a fundamental role in these issues, reflected in the following hashtags: #teachers #teachers #education #primary #primary #neuroeducation #students #school #teacher #libraries.

In relation to the professions of the most viral profiles, (i) the one that receives the highest number of likes is linked to digital creators [T1], which is logical, as their aim is to achieve impact on social media. Unlike an association that is hidden in these social media whose purpose is to raise awareness and collaborate with the cause that moves them [T34], but given the neutral messages and content they publish, they are relegated to the background due to a "glass algorithm" in social networks, i.e. they are visible to any user who explicitly searches for them, but not by suggestions from digital platforms.

4. Discussion and conclusions

Influencing groups plays a relevant role as actors who are phagocytising the influence of other actors with more altruistic and humanist interests. This is an important issue if we take into account that the issues analysed affect groups and people who are often unequal, and who need to make their demands, experiences and achievements visible. Social networks as spaces in which perception and reality are intermediated, with effects on the ways of understanding it and what is or is not important for the collective public debate, plays in favour of groups and actors with political interests (Del-Fresno-García, 2019). In this sense, the role of social media as spaces for affinity, participation and collaboration, as evidenced by the work of Grace et al. (2019) and De-Groot et al. (2022), has not been confirmed in this study and for the subject under analysis. The role of social media as opportunities for users to interact and share information in influential relational structures of support and monitoring has not been validated in this research either.

The only way to reverse the viral power of influencers is twofold: to change in the algorithms of digital platforms, and make users on social media aware. For the first possibility, it would require state and supra-state public policies that promote some law to establish some control over influencers, with the paradox, as demonstrated in this study, that they are the most influential actors in the networks. For the second possibility, a digitally competent teaching staff is required. The results obtained point to the need to raise awareness among digital citizens with a critical vision in order to understand how algorithms impose the messages of certain actors, as opposed to other groups interested in making situations related to disability visible, and to claim their educational and social value from inclusive references. Some of the 100 most viral Instagram hashtags (#) confirm these interests. Teachers must make new generations understand the asymmetry in communication, as represented by Barberá et al. (2015) and Brady et al. (2019). This will enable them to recognise that certain publications, and the profiles associated with them, define invaluable human and social values.

In relation to this, it is important that teachers help the new generations to become progressively aware of the risks of falling into bubble filters or echo chambers. This would avoid a dominance as a factor of destabilisation of democracies worldwide, in the face of disinformation promoted by influential groups according to their interests, which defines another important challenge for teachers (Ortega-Sánchez et al., 2021). In reference to the study Robles et al. (2022) on bots and negative feelings associated with them as fuel for political polarisation, it has no presence in the social and educational subject matter under analysis. This may be due to public awareness, which would not assume and accept explicit confrontation, as is the case with topics such as immigration, the economy, security, politics or health. These topics are prone to bots that polarise publications, and thus, citizens. The relegation of the power of influence of influential groups on international days would support this conclusion. Other results that would validate these conclusions are that, although the male gender has more followers than the female gender, it is women who get more likes. This could be linked to the positive feelings that prevail in their publications for the content analysed. These results would indicate another important task for teachers: teaching new generations to post on social media with constructive attitudes, away from negative emotions (Arcila-Calderón, 2022).

The innovative aspect of the use of the Business Intelligence tool justifies its consideration as content of the digital competence of teachers for education in the coming years. The scientific-technical impact implies a radical change in the way of conceptualising and using social media as spaces for communication and construction of discourses and actions in the field of inclusive education and disability. Among other reasons, this occurs because our analysis no longer views social media as a tool for the uncritical and unreflective consumption of discourses and superficial contacts, but as valuable media that support relationships and content of educational and social value. Another important fact that supports the innovative nature of the tools and techniques applied lies in the case studies we have initiated, because they will allow us to systematise the value and function of multimodal representation to communicate and promote awareness, debates, and narratives that generate counter-hegemonic actions. Another reason lies in the methodological impact inherent to the study, defined by the analytical techniques we will carry out. In the scientific debate on communication and human thought, this will contribute to define, the value of interdisciplinary work in two fields, that of Educational Technology and that of Telecommunications.

Finally, it should be noted that the analysis process carried out with Big Data and Business Intelligence tools is in itself a digital competence in inclusive relationship spaces in educational institutions, which will help teachers to understand the relational structure of social networks and to know how to arrive at meanings about the content of the exchanges made by the participants in these networks. This implies that such teaching competence includes knowing how to identify the most influential profiles and how to investigate the meaning given to topics and behaviours, which are not always visible (Del-Fresno-García, 2014), linking them to messages on social media and the type of profiles associated with them. This knowledge will lead teachers to acquire a collective ethic, for which it will sometimes be necessary to unlearn ideas and beliefs built in the history of each of the participants in their life contexts. This is the only way to build a shared morality based on reason, and not only on emotions, in the contemporary, mixed and diverse relationships that characterise educational centres that seek inclusive teaching situations. This requirement in teaching practice entails the need to assume the inquiring-innovative dimension of social networks, of a critical nature, as one of the objectives of teachers' digital competence (Bautista, 2021).

Authors' Contribution

Idea, C.B.; Literature review (state of the art), C.B., L.R; Methodology, C.B; Data analysis, C.B. L.R; Results, C.B, L.R, A.B; Discussion and conclusions, C.B, L.R, A.B.; Drafting (original draft), C.B, L.R, A.B.; Final revisions, C.B, L.R, A.B; Project design and sponsorships, C.B, L.R, A.B.

Funding Agency

This work has been supported by the Madrid Government (Comunidad de Madrid-Spain) under the Multiannual Agreement with Universidad Complutense de Madrid in the line Excellence Programme for university teaching staff, in the context of the V PRICIT (Regional Programme of Research and Technological Innovation).

References

Arcila-Calderón, C., Sánchez-Holgado, P., Quintana-Moreno, C., Amores, J., & Blanco-Herrero, D. (2022). Hate speech and social acceptance of migrants in Europe: Analysis of tweets with geolocation. [Discurso de odio y aceptación social hacia migrantes en Europa: Análisis de tuits con geolocalización]. *Comunicar*, *71*, 21-35. https://doi.org/10.3916/C71-2022-02

Awidi, I.T., Paynter, M., & Vujosevic, T. (2019). Facebook group in the learning design of a higher education course: An analysis of factors influencing positive learning experience for students. *Computers & Education*, *129*, 106-121. https://doi.org/10.1016/j.compedu.2018.10.018

Barberá, P. (2015). Birds of the same feather tweet together: Bayesian ideal point estimation using twitter data. *Political Analysis*, *23*(1), 76-91. https://doi.org/10.1093/pan/mpu011

Barberá, P., Jost, J.T., Nagler, J., Tucker, J.A., & Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science*, *26*(10), 1531-1542. https://doi.org/10.1177/0956797615594620

Bautista, A. (2021). Functional resignification and technological innovation as a digital teaching competence. *IEEE. Revista Iberoamericana de Tecnologías del Aprendizaje*, *16*(1), 93-99. https://doi.org/10.1109/RITA.2021.3052679

Becker, L.T., & Gould, E.M. (2019). Microsoft power BI: Extending excel to manipulate, analyze, and visualize diverse data. *Serials Review*, *45*(3), 184-188. https://doi.org/10.1080/00987913.2019.1644891

Brady, W.J., Wills, J.A., Burkart, D., Jost, J.T., & Van-Bavel, J.J. (2019). An ideological asymmetry in the diffusion of moralized content on social media among political leaders. *Journal of Experimental Psychology: General*, *148*(10), 1802-1813. https://doi.org/10.1037/xge0000532

Brunner, M., Palmer, S., Togher, L., & Hemsley, B. (2019). I kind of figured it out': The views and experiences of people with traumatic brain injury in using social media-self-determination for participation and inclusion online. *International Journal Language & Communication Disorders*, *54*(2), 221-233. https://doi.org/10.1111/1460-6984.12405

Carlsson (Ed.) (2019). *Understanding media and information literacy (MIL) in the digital age. A question of democracy*. University of Gothenburg. https://bit.ly/3xWsCz2

Cheng, L.C., & Tsai, S.L. (2019). Deep learning for automated sentiment analysis of social media. In *Proceedings of the 2019 IEEE/ACM international conference on advances in social networks analysis and mining* (pp. 1001-1004). Association for Computing Machinery. https://doi.org/10.1145/3341161.3344821

De-Groot, R., Kaal, L.H., & Wouter, P.S. (2022). The online lives of adolescents with mild or borderline intellectual disabilities in the Netherlands: Care staff knowledge and perceptions. *Journal of Intellectual & Developmental Disability*. https://doi.org/10.3109/13668250.2021.2004635

Del-Fresno-García, M. (2014). Haciendo visible lo invisible: Visualización de la estructura de las relaciones en red en Twitter por medio del análisis de redes sociales. *Profesional de la Información*, *23*, 246-252. https://doi.org/10.3145/epi.2014.may.04

Del-Fresno-García, M. (2019). Desórdenes informativos: sobreexpuestos e infrainformados en la era de la posverdad. *Profesional de la Información*, *28*(3), 1-11. https://doi.org/10.3145/epi.2019.may.02

Dellwing, M. (2021). *The continuities of Twitter strategies and algorithmic terror. Society for the study of symbolic interaction*. Sage. https://doi.org/10.1002/symb.585

Ditrendia (Ed.) (2009). *Mobile en España y en el Mundo*. Ditrendia Digital Marketing Trends. https://bit.ly/3MNim3j

Grace, E., Raghavendra, P., Mcmillan, J., & Gunson, J. (2019). Exploring participation experiences of youth who use AAC in social media settings: Impact of an e-mentoring intervention. *Augmentative and Alternative Communication*, *35*(2), 132-141. https://doi.org/10.1080/07434618.2018.1557250

Hemsley, B., Dann, D., Palmer, S., Allan, M., & Balandin, S. (2018). Using Twitter to access the human right of communication for people who use Augmentative and Alternative Communication. *International Journal of Speech-Language Pathology*, *20*(1), 50-58. https://doi.org/10.1080/17549507.2017.1413137

Ladogina, A., Samoylenko, I., Golovina, V., Razina, N., & Petushkova, E. (2020). Communication effectiveness in social networks of leading universities. *Diálogo*, *43*, 35-50. https://doi.org/10.18316/dialogo.v0i43.6497

Liu, B. (2010). Sentiment analysis and subjectivity. In N. Indurkhya, & F. J. Damerau (Eds.), *Handbook of Natural Language Processing*. Chapman and Hall/CRC. https://doi.org/10.1201/9781420085938

Mitchell, R. (2018). *Web scraping with Python: Collecting more data from the modern web*. O'Reilly Media: O'Reilly Media, Inc. https://bit.ly/3zw5nOn

Ortega-Sánchez, D., Blanch, J.P., Quintana, J.I., Cal, E., Fuente-Anuncibay, C., & Hate, R. (2021). Speech, Emotions, and Gender Identities: A Study of Social Narratives on Twitter with Trainee Teachers. *International Journal of Environmental Research and Public Health*, *18*(8), 4055-4055. https://doi.org/10.3390/ijerph18084055

Robles, J., Guevara, J., Casas-Mas, B., & Gómez, D. (2022). When negativity is the fuel. Bots and Political Polarization in the COVID-19 debate. [Cuando la negatividad es el combustible. Bots y polarización política en el debate sobre el COVID-19]. *Comunicar*, *71*, 63-75. https://doi.org/10.3916/C71-2022-05

Tuzel, S., & Hobbs, R. (2017). The use of social media and popular culture to advance cross-cultural understanding. [El uso de las redes sociales y la cultura popular para una mejor comprensión intercultural]. *Comunicar*, *51*, 63-72. https://doi.org/10.3916/C51-2017-06

We are social (Ed.) (2021). *Digital 2021. The latest insights into the 'state of digital'*. We Are Social Ltd. https://bit.ly/3POIhJV

White, P., & Forrester-Jones, R. (2020). Valuing e-inclusion: Social media and the social networks of adolescents with intellectual disability. *Journal of Intellectual Disabilities*, *24*(3), 381-397. https://doi.org/10.1177/1744629518821240

Comunicar, n. 74, v. XXXI, 2023 | Media Education Research Journal | ISSN: 1134-3478; e-ISSN: 1988-3478
www.comunicarjournal.com

(Un)founded fear towards the algorithm: YouTube recommendations and polarisation

Miedo (in)fundado al algoritmo: Las recomendaciones de YouTube y la polarización

Dr. Javier García-Marín. Professor, Department of Political Science and Administration, University of Granada (Spain) (jgmarin@ugr.es) (https://orcid.org/0000-0002-2766-0266)
Dr. Ignacio-Jesús Serrano-Contreras. Associate Researcher, SINAI Research Group, University of Jaen (Spain) (ijserran@ujaen.es) (https://orcid.org/0000-0002-2399-0647)

ABSTRACT

Social media have established a new way of communicating and understanding social relationships. At the same time, there are downsides, especially, their use of algorithms that have been built and developed under their umbrella and their potential to alter public opinion. This paper tries to analyse the YouTube recommendation system from the perspectives of reverse engineering and semantic mining. The first result is that, contrary to expectations, the issues do not tend to be extreme from the point of view of polarisation in all cases. Next, and through the study of the selected themes, the results do not offer a clear answer to the proposed hypotheses, since, as has been shown in similar works, the factors that shape the recommendation system are very diverse. In fact, results show that polarising content does not behave in the same way for all the topics analysed, which may indicate the existence of moderators –or corporate actions– that alter the relationship between the variables. Another contribution is the confirmation that we are dealing with non-linear, but potentially systematic, processes. Nevertheless, the present work opens the door to further academic research on the topic to clarify the unknowns about the role of these algorithms in our societies.

RESUMEN

Las redes sociales han instaurado una nueva forma de comunicarse y entender las relaciones sociales. A su vez, en lo que podría entenderse como un aspecto negativo, los algoritmos se han construido y desarrollado bajo el paraguas de un amplio abanico de conjeturas y diferentes posiciones al respecto de su capacidad para dirigir y orquestar la opinión pública. El presente trabajo aborda, desde los procesos de ingeniería inversa y de minado semántico, el análisis del sistema de recomendación de YouTube. De este modo, y, en primer lugar, reseñar un resultado clave, las temáticas analizadas de partida no tienden a extremarse. Seguidamente, y mediante el estudio de los temas seleccionados, los resultados no ofrecen una clara resolución de las hipótesis propuestas, ya que, como se ha mostrado en trabajos parecidos, los factores que dan forma al sistema de recomendación son variados y de muy diversa índole. De hecho, los resultados muestran cómo el contenido polarizante no es igual para todos los temas analizados, lo que puede indicar la existencia de moderadores –o acciones por parte de la compañía– que alteran la relación entre las variables. Con todo ello, trabajos como el presente abren la puerta a posteriores incursiones académicas en las que trazar sistematizaciones no lineales y con las que, tal vez, poder arrojar un sustento más neto y sustancial que permita despejar por completo parte de las dudas sobre el papel de los algoritmos y su papel en fenómenos sociales recientes.

KEYWORDS | PALABRAS CLAVE

Machine learning, YouTube, social media, recommendation system, polarisation, communication.
Aprendizaje de máquina, YouTube, redes sociales, sistemas de recomendación, polarización, comunicación.

Received: 2022-05-30 | Reviewed: 2022-06-21 | Accepted: 2022-07-13 | OnlineFirst: 2022-10-30 | Published: 2023-01-01
DOI https://doi.org/10.3916/C74-2023-05 | Pages: 57-66

1. Introduction and theoretical framework

The rise of ICTs has produced new scenarios both for social interaction and for the consumption of information and entertainment in societies at the end of the 20th century, especially in the first decades of the 21st century. In addition to the well-known mass media (including the press, radio, and television), other methods of disseminating and accessing messages on a mass scale have been added, which has also generated new symbioses in which the roles of the communication chain have tended to become blurred and hybridised (Berrocal-Gonzalo et al., 2014). Finally, we talk about the benefits and contributions of the Internet to democratise information dissemination mechanisms (Arias-Maldonado, 2016; Nielsen & Fletcher, 2020). While the Internet and its satellites were growing (boosted by the development of computing and technological advances in processing), so were the multiple studies on its effects. In this respect, this (inter)connected web that different authors (Berners-Lee, 2000), and with different adjectives (see McLuhan (1959), Habermas (1981) or Castells (2001) as a summary of these expressions), has engendered first the Web 1.0 and, later, the 2.0 (O'Reilly & Battelle, 2009), as well as other concepts (Latorre, 2022).

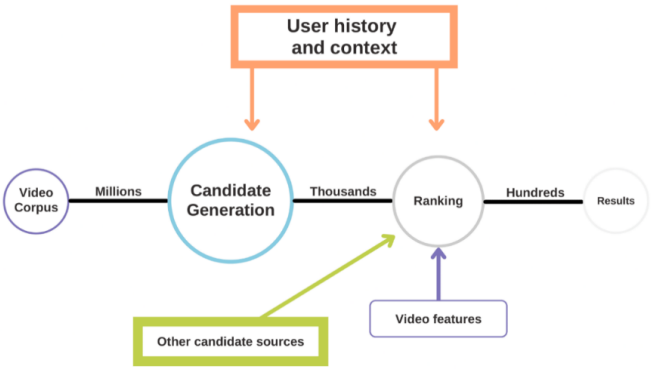
These developments give rise to various reflections on the effects they may have on citizens. Some of these new reflections already pointed to how the rise of the growing Web 2.0 could become the axis of a paradigm shift, ultimately posing a challenge and opportunity for both political spheres and liberal democracies in the 21st century (Sunstein, 2007; Lilleker & Jackson, 2008; Chadwick, 2009; Howard, 2021; Messina, 2022). However, the significant advances of the interconnected world were not yet fully consolidated. The emergence of what has come to be known as social media has led to a reconfiguration of the development of human relationships (Wigand et al., 2010). Thus, incorporating Facebook, Twitter, Instagram, or YouTube into everyday life has meant a change for contemporary societies, with users from all corners of the planet, as well as having a reach and diffusion ratio of more than a third of the world's population. Now that we are immersed in this transition period, the focus is on the effects that may arise from this new social and media drift driven by the networks. In this respect, part of the research carried out in academia and other spheres, such as journalism or politics, has focused on the internal side of interconnection. In this case, we are talking about the role that algorithms, especially their architectures and protocols, can play as mediators of the communication process.

Now, those related questions arise with the perspective of expanding algorithms. In this context, approaches emerge that point to the possible relationship between the role of social networks - although with special emphasis on their computational models - and the search for an understanding of singular social events: from a greater presence of political polarisation in public debate (Hernández et al., 2021), to situations that are complex to define, such as Brexit or the victory of Donald Trump in 2016. Despite the clear correlations that may exist between one event and another, the truth is that the research points to different sides without reaching conclusions. At least this is established in the doubts presented by works such as those of Rasmussen and Petersen (2022), Bail (2021) or Barberá (2020), who point to multifactorial, and even those who point to the analogical plane (Arceneaux & Johnson, 2010) as a key axis to reach an answer. This situation is therefore complex and subject to ambivalent dynamics. Based on this, the present research delves into some of the phenomena behind YouTube and its capacity to flood the multiple spheres of the media scene (Banaji, 2013). As Yesilada and Lewandowsky (2022) also point out, one of the critical factors focuses on the complexity of understanding its system. In this way, and as already pointed out by other studies, such as those of Luengo et al. (2021) or Serrano-Contreras et al. (2020), this paper aims to point out the drifts that the algorithm can generate. In addition, it seeks to incorporate into the debate whether this computational model is of any use in considering the emergence of social phenomena such as polarisation (Van-Bavel et al., 2021).

2. Data and method

This research proposes an analysis of the YouTube algorithm from multiple perspectives. In this regard, under the so-called reverse engineering process (Rekoff, 1985), as well as using text mining techniques and semantic measurement indexes, we seek to shape a progressive understanding of what lies beneath the computational architectures implemented by YouTube.

Figure 1. Neuronal model: Candidate generation



Note. Covington et al. (2016).

To this end, we seek to leave aside some of the main functions that Alphabet, as the owner of YouTube, incorporates into the training and subsequent development of the algorithm of its video server (see the work of Alphabet's employees, Davidson et al. (2010) or Covington et al. (2016), as well as Figure 1, for an approximation of the model that the platform uses to offer the results to the user). In this way, the goal is to try to parameterise the behaviour of the model avoiding the set of passive and active data that we provide on the network while browsing (e.g. location; search history; personal data...), since all these metrics are used to compile detailed information about our supposed interests - fundamental to our usage experience (Dimopoulos et al., 2013); On the other hand, we seek to review the model. These experiences are the basis for assumptions which are indispensable for understanding some of the most commonly used nomenclatures when discussing algorithms, such as the bubble filter (Pariser, 2017) and other externalities (Bishop, 2018).

Within this set of effects that the algorithmic era can causethere is a factor that has become very popular recently, the idea of radicalisation. To support these ideas about the reinforcement of a position, we used work, research, and empirical examples, which have shown how the platform's algorithm tended to become more and more extreme (Tufekci, 2018; Alfano et al., 2021; Almagro & Villanueva, 2021; Chen et al., 2021).

Therefore, the behaviour of YouTube's recommendation algorithm, based on users' interests, should lead to higher consumption of related materials. In other words, it will interpret users' searches as interests and thus try to make it easier for them when searching, by recommending similar videos (which would result in a filter bubble). It follows that recommendations could result in greater polarisation by causing less exposure to different viewpoints or topics. Hence, the first hypothesis we propose in this research is:

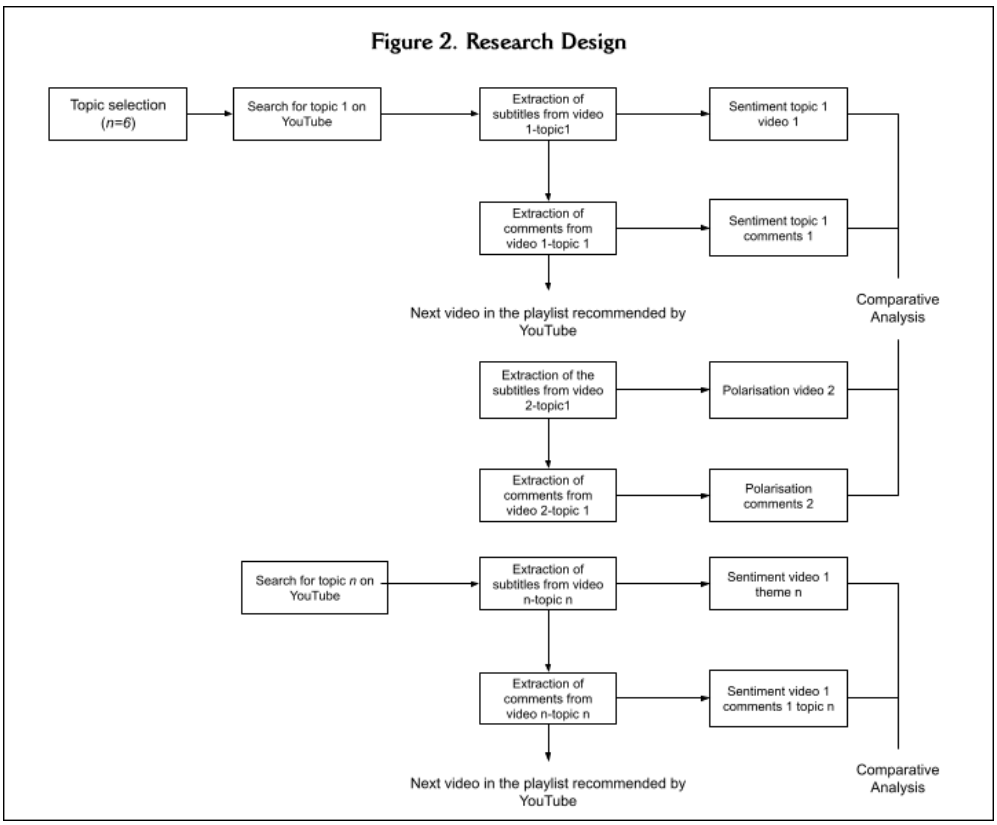
- H1: Videos recommended by YouTube will become increasingly extreme or polarised.

At the same time, such recommendations could create communities with very similar interests (called homophilic classifications), which, in turn, could be related to echo chambers. Therefore, our second research hypothesis is:

- H2: Comments on videos recommended by YouTube will be equally polarised.

The logic of the research is that, according to part of the scientific community[1], YouTube's algorithm recommendations move towards extrapolating the interests raised by users' searches, creating this filter bubble and, at the same time, increasing the polarisation of the content shown (by being increasingly focused on very specific content or a specific point of view). Users should thus behave similarly.

Figure 2 describes the research we have designed to test the hypotheses. First, we selected an existing YouTube account belonging to one of the researchers, which had never been used (we could say that there was no initial metadata linked to the account, so the searches we initiated would create the metadata about our interests)[2]. Subsequently, we chose topics on which to apply the design. We tried to ensure relevant topics that served to discriminate specific moderators that could affect the relationship between the variables we intended to analyse.



Thus, the topics[3] were:

- (Spanish) National politics: search for "política nacional."
- (Spanish) Political parties at ideological extremes: searches for "Podemos" and "Vox."
- Vegetarianism: search for "comida vegetariana."
- Conflict: search for "guerra en Ucrania."
- Feminism: search for "feminismo."
- COVID: search for "COVID."

In other words, topics differed in conflict and were related to various fields, from conflict to vegetarianism. All were sensitive topics where we expected significant comments (on the videos that allowed it).

The next step was to start searching for the topic on YouTube, both on the proposed account and without any account, and analyse the first video that the results recommended. When the video was played, we continued with the automatic playback. Thus, we could see the recommendation made by YouTube's algorithm and up to a certain number of videos (with a maximum of 100 per topic). So, we have at least two first videos per topic, one searching with the account and the other without. The goal was to see whether the algorithm behaved differently in the absence of account metadata.

The videos were analysed by extracting their content through the subtitles (therefore, it is an analysis of the textual script, not the images). As this is a resource that is not present in all the videos on the platform, it was decided to apply the analysis to one video out of every 10 in the playlist (or the one that was closest in ordinal order), provided that there was a sufficient number of videos (as will be explained, in the case of the videos captured without an account on the platform, the text was extracted from all the videos that contained it). With this text, we proceeded to an affective polarisation analysis (on the debate around the concept, see Iyengar et al., 2019) using a technique that we had used on previous occasions with considerable success (see Serrano-Contreras et al., 2020, for a detailed explanation). The procedure consisted of modifying a sentiment analysis (the selected tool was Orange3, Demsar et al., 2013, based on Python and using a multilingual dictionary for more than 50 languages). The modification consisted of

calculating the mean sentiment of a given dataset and measuring the distance between the unit of analysis and the overall sample (thus being an affective distance analysis). In this case, we varied the analysis to measure the polarisation of videos and comments based on the sentiment of the first video by topic and account. Thus, the polarisation of the other videos and comments is given as the distance from the first one, but in absolute numbers. This way, we can tell if a particular video has a tone that is distant (positive or negative) from the first one. We do not look at how negative or positive the videos are, as these are circumstantial considerations and would require a detailed analysis of the content, but only at their distance from the first one. We understand that affective polarisation can occur through positive and negative charges (for example, cheering or congratulating a terrorist group). The same analysis was done concerning comments, except all possible comments were analysed, including videos without subtitles (although not all videos allow comments). The numerical results range from 0 to 100, but it is common to obtain very low numbers (around 0.1-2). This is because most of the content is not affectively loaded (even after pre-processing the text with the usual techniques, as was done). The only consequence is that minor changes in value indicate substantial variations in affective polarisation; after all, we are talking about millions of analysed words (Table 2).

For each video, then, the following variables were extracted: those provided by the platform (number of likes, views, number of comments, etc.), topic, position in the automatic playlist (platform recommendation), polarisation of the video's content (one out of ten in those extracted with an account) and polarisation of the video's comments (in those that have them enabled). Finally, tables 1 and 2 describe the number of units of analysis (750 videos and, including comments, nearly three million words).

Table 1. Analysed videos. By account and topic

	Account	No account	Total
Vegetarian food	101	6	107
COVID	100	5	105
Feminism	101	11	112
War in Ukraine	103	7	110
Podemos	89	4	93
National politics	108	0	108
VOX	109	6	115
Total	**711**	**39**	**750**

Note. The results of the polarisation analysis on videos and comments are then analysed in a comparative way. As the hypotheses indicate, each video, and its comments in the playlist is expected to be slightly more polarised than the previous one.

3. Findings

One element that needs to be highlighted is that YouTube changes its parameters and ways over time. Therefore, in this research, we could not work with both "likes" and "dislikes," as only positive data is now provided. However, we do not believe that this will affect the research.

Table 2. Number of analysed words. By account, topic and word origin

	Account		No Account		Total
	Subtitles	Comments	Subtitles	Comments	
Vegetarian food	55,394	323,179	17,634	31,947	428,154
COVID	57,378	200,049	19,513	24,526	301,466
Feminism	35,514	361,534	31,337	55,080	483,465
War in Ukraine	14,989	221,559	2,324	16,327	255,199
Podemos	55,748	251,270	14,005	12,263	333,286
National politics	58,693	405,947	0*	0*	464,640
VOX	43,011	445,720	15,598	14,542	518,871
Total	**320,727**	**2,209,258**	**100,411**	**154,685**	**2,785,081**

Note. In the case of national policy, there were no consistent results, probably due to the lack of metadata on the country of origin (the Tor network was used to avoid them).

Along the same lines, YouTube does not provide the exact same content to one user as to another, something that is evident from what has already been outlined by Pariser (2017) and from daily consumption. However, there is another fact to take into consideration. Logging in with or without an account produces different results (Table 1). Despite the obvious, there is another interesting element that has been found in this work: when an automatic playback is carried out, if the process is conducted with an account, the model continues to offer videos, but if it is done without an account, the model ends

up entering a loop in which two videos tend to playrepeatedly. Hence, the need to place a limit on the 100th video is unnecessary in the case of access without an account. This also affects the length of the videos. While the length of the videos did not seem to be a factor without an account and have set a constant consumption, the model tended to offer longer and longer videos.

Another fact to consider, and already noted in the evolution from Davidson et al. (2010) to Covington et al. (2016), is that YouTube's ranking of results varies over time. This seems to be evidenced by the type of content offered by autoplay. This is seen in both the topic of feminism and vegetarianism. While the former seems to be influenced by the fact that a video from the TED talks channel was selected in the first instance, in the latter, even though it could generate opposing positions such as vegetarianism, the type of video, in this case, based simply on cooking recipes, has meant that the algorithm has not tended towards other paths as has happened with the rest of the topics, which bifurcated and diverted to other areas. For example, in the case of the conflict in Ukraine, a large part of the final sample is made up of relaxing music videos. Thus, there are factors that the algorithm aims to reward in order to filter a certain content to offer the user. This position seems to be a clear commitment by the company (see also Goodrow (2021) and Mohan (2022) for a detailed explanation of the actions undertaken by the platform to create content that is less harmful for both information and consumption by users of all age ranges). Hence, Table 2 shows data that can sometimes be paradoxical, such as the fact that there are more words analysed from subtitles than from comments.

Table 3. Mean polarisation

	Account		No Account	
	Subtitles	**Comments**	**Subtitles**	**Comments**
Vegetarian food	0.23	0.81	0.11	0.14
COVID	1.84	2.13	0.18	1.00
Feminism	0.67	0.96	0.45	0.48
War in Ukraine	0.74	0.99	0.03	0.49
Podemos	0.21	1.29	0.23	0.11
National politics	0.56	0.30	-	-
VOX	0.96	0.41	0.21	2.16
$\bar{X}$	**0.76**	**0.91**	**0.22**	**0.67**

Let us look at the average aggregate polarisation data (Table 3). It is easy to observe the different values depending on the topics, the accounts, and whether they come from the videos themselves or the comments. In the first case, there clearly are topics where polarisation is higher, especially those referring to "COVID" ($\overline{X}$=1.35), followed by "VOX" (0.68), "national politics" (0.56) and "feminism" (0.55). Although results are not surprising, there are essential differences between those analysed with and without an account (especially in the case of "VOX" and "COVID").

Table 4. Correlation Matrix (Pearson)

	ViewCount	Comments	Likes	Order	VisitCount	Pol-Com	Pol-Sub
ViewCount	1	.679**	.808**	-.069	-.113	-.064	.161
N	755	755	755	755	711	676	113
Comments	.679**	1	.786**	-.059	-.062	-.174	.045
N	755	755	755	755	711	676	113
Likes	.808**	.786**	1	-.027	-.053	-.146	-.076
N	755	755	755	755	711	676	113
Order	-.069	-.059	-.027	1	.132**	.124**	.316**
N	755	755	755	755	711	676	113
VisitCount	-.113**	-.062	-.053	.132**	1	-.082*	-.060
N	711	711	711	711	711	632	69
Pol-Com	-.064	-.174**	-.146**	.124**	-.082*	1	.265**
N	676	676	676	676	632	676	113
Pol-Sub	.161	.045	-.076	.316**	-.060	.265**	1
N	113	113	113	113	69	113	113

Note. *Correlation is significant p=<0.01 (2-tailed). **Correlation is significant p=<0.05 (2-tailed).

Except for "COVID" (2.07), user comments do not show a similar pattern, with "war in Ukraine" (0.95), "Podemos" (0.95) and "feminism" (0.88), in addition to the aforementioned "COVID," being the topics where the most significant polarisation has been observed. Although there are also notable differences concerning the origin of the video (with or without an account), the data seem to point to the existence of a possible echo chamber in some cases or, at least, to a certain degree of agreement between

users with an account who comment on the videos we have catalogued as "VOX" and "national politics," while those who comment on the videos captured with the searches "COVID," "Podemos" or "Ukrainian war," show signs of highly polarised comments. Overall, the videos analysed with a Google account have an average polarisation of 0.76 with a fairly high dispersion (+- 0.70), while those analysed without it have a much lower polarisation of 0.22 (+-0.25). The distance is smaller if we include the polarisation in the users' comments, 0.91 and 0.67 respectively; in both cases, it is substantially higher than that obtained from the videos.

As can be seen in Table 4, the correlation matrix provides exciting data. On the one hand, we have the expected correlations, such as the relationship between likes, comments, and views; the correlations are robust because, in essence, they are measuring the same thing: the popularity of a given video. On the other hand, the variable we are most interested in is "Order," which indicates the position in the playlist offered by YouTube, and this is where there are interesting findings. Fundamentally, we can see a positive correlation between polarisation and the order of the video, both in the comments (Pol-Com) and in the video itself (Pol-Sub) and between themselves. Indeed, this is not an excessively strong relationship, but it is not negligible either, especially the relationship between the polarisation of the video and the order (.316).

With regards to video polarisation, it is interesting to explore this more closely. The relationship between the two variables is not linear (the results of attempting to model it using linear regression have been unsuccessful, x2=-0.002), so it is likely there is a moderator. It was decided to analyse the means to discover the variable that may be altering this relationship and to recode the Order variable into three segments (Order-Cat): one to three videos, four to nine, and more than nine. The segmentation is, of course, not arbitrary. We have estimated that it is possible and even likely, that a user will watch up to three videos proposed by YouTube in a row. We find it less likely that between four and nine videos will be watched and quite unlikely that more than nine videos in a row suggested by the algorithm will be watched. Thus, we consider that, in the first case, we would face a low exposure to the algorithm, medium in the second, and high in the third.

Table 5. Mean comparison (only the significant ones are displayed)

	Order-Cat		
	(low exp.)	(medium exp.)	(high exp.)
	Mean	Mean	Mean
Veg-Food (Pol-Com)	0.11	0.26	0.86
Covid (Pol-Com)	0.41	1.65	-
Covid (Pol-Sub)	0.25	0.19	2.15
Feminism (Pol-Com)	0.26	0.70	-
Ukraine (Pol-Com)	0.32	0.29	1.07
Nac-Pol (Pol-Sub)	0.20	-	0.63
Podemos (Pol-Com)	0.07	-	1.61
VOX (Pol-Com)	-	1.26	0.42
VOX (Pol-Sub)	0.12	0.26	1.07

Note. Results are based on two-tailed tests assuming equal variances. Tests are adjusted for all pairwise comparisons using the Bonferroni correction. All significance levels are .05.

Table 5 shows the mean comparison results, although only in those cases where the test indicates significant differences. The first impression is that each topic seems to behave differently: the videos (Pol-Sub) offered by the platform when searching for the terms "COVID", "national politics", and "VOX", indeed tend to be more polarised in each of the three proposed sections (except the first to the second of "COVID"). The case of "VOX" is perhaps the clearest and it shows how the average polarisation increases in each section in an almost linear fashion. In regard to comments, the topics where differences can be seen are "vegetarian food," "COVID," "feminism," "Ukrainian war," "Podemos," and "VOX." In other words, on all topics. However, unlike the case of the videos, we observe different trends: while in all of them, the tendency continues to be towards an increase in polarisation, the opposite is observed in VOX.

In the case of both the videos and the comments, we expected a similar relationship between the different topics. However, the differences are so substantial that it is challenging to validate the initial hypotheses. There does appear to be some relationship between the position of the video - and its

comments - and an increase in polarisation, but the exceptions prevent us from claiming a direct relationship. Perhaps some particularly prominent cases are being modified by the company itself, as with COVID, where an effort was made to ensure that citizens received less extreme information. In any case, results indicate moderators that are difficult to measure in this relationship.

4. Discussion and conclusions

Approaches to the debate on the role of algorithms are slowly taking hold. Despite this, although the various positions are beginning to make clear outlines of what we may face, other aspects are still challenging to address. This is the case with recommendation systems (Yesilada & Lewandowsky, 2022). In this particular area, works such as the present one, hopefully, serve to portray what happens in processes such as those of YouTube. At least through this sampling model, it seems evident that the chosen themes do not tend to become extreme. In other words, playing a video about vegetarian food does not end up with a video about veganism or "anti-speciesist" movements. That circumstance seems to be exposed to more complex factors than the simple issue of search, not least the ability of YouTube's training model to access user data and interests and generate a consumption pattern. This, in turn, it is a variable to be considered when carrying out data collection actions with this series of sampling methods. However, we cannot give a conclusive answer. As shown at the beginning, theorists state that recommendation algorithms may be causing the phenomena called filter bubbles and echo chambers (Terren & Borge-Bravo, 2021). However, the results presented here are not conclusive: there are topics where the algorithm does seem to behave in that direction, but there are other topics where such a relationship, at least with our research strategy, is not perceived.

Nevertheless, the results offered are striking and consistent with previous research, where we found that the topic strongly moderates content polarisation and user reactions (Serrano-Contreras et al., 2020). This opens up other questions worthy of further research: to find whether topics are a moderator in this relationship. For these topics, could we claim that these hypotheses hold? And, more importantly, why? Unfortunately, the data offered here are insufficient to answer these questions beyond the description of those chosen. However, they humbly contribute to the direction that the relationship, if it exists, is neither direct nor linear. Nevertheless, the results may be different with a selection based on other groupings, e.g., topics on politics, extremism, music, etc., as well as implementing coherent monitoring of clusters of so-called prosumers. Moreover, with the proper internal coherence, conclusions that move towards the definitive answer to the question that theorists have been asking for a decade or more could be reached: do social networks polarise our citizens?

Notes

[1] It should be noted that part of the actions carried out to obtain the samples have been based on the methodological recapitulation, in one way or another, of previous empirical work, which systematically continued consumption in the case of automatic reproduction. This clarification is done mainly because the authors consider that this type of media consumption is very different from most of the actions that users carry out on video platforms -mainly of short content. Therefore, this type of action, based on constant consumption without pauses or alterations in the periodicity of consumption, directly affects the results that the algorithmic model will end up offering.

[2] The present work has used several incursions to obtain the sample data. Before the detailed explanations, it should be pointed out that this type of analysis can also be carried out from the API to access its servers. However, it requires a login account, which was discarded as it was not considered an organic process for the collection. Firstly, we conducted a search using automatic video playback without having a linked account and rejected all factors that could feed microtargeting. On the other hand, given the limited amount of data obtained, we resorted to anonymisation techniques by Onion using layers through Tor's incognito model via Brave. In addition, the VPN provided by the University of Granada was used to add another layer. Having carried out the same process, we found the same dilemma as the previous search: the lack of magnitude in the sample and cessation of activity due to the lack of interaction with the platform. Therefore, in the end, as mentioned above, we undertook the process through a user account with no activity.

[3] All keyword searches were conducted in lower case and with the appropriate Spanish accents.

Authors' Contribution

Idea, J.G.M., I.J.S.C.; Literature review (state of the art), J.G.M., I.J.S.C.; Methodology, J.G.M., I.J.S.C.; Data analysis, J.G.M., I.J.S.C.; Results, J.G.M., I.J.S.C.; Discussion and conclusions, J.G.M., I.J.S.C.; Drafting (original draft), J.G.M., I.J.S.C.; Final revisions, J.G.M., I.J.S.C.; Project design and sponsorships, J.G.M., I.J.S.C.

Funding Agency

This research is part of the R&D&I project PID2021-128272NB-I00 funded by the Spanish Ministry of Science and Innovation MCIN/ AEI/10.13039/501100011033/ ERDF "A way of doing Europe".

References

Alfano, M., Fard, A.E., Carter, J.A., Clutton, P., & Klein, C. (2021). Technologically scaffolded atypical cognition: The case of YouTube's recommender system. *Synthese*, *199*(1-2), 835-858. https://doi.org/10.1007/s11229-020-02724-x

Almagro, M., & Villanueva, N. (2021). Polarización y tecnologías de la Información: Radicales vs. extremistas. *Dilemata*, *34*, 51-69. https://bit.ly/38YwIiH

Arceneaux, K., & Johnson, M. (2010). Does media fragmentation produce mass polarization? Selective exposure and a new era of minimal effects. In A. Campbell, & L. Martin (Eds.), *American Political Science Association 2010 Annual Meeting*. SSRN. https://bit.ly/3M1e7jJ

Arias-Maldonado, M. (2016). La digitalización de la conversación pública: Redes sociales, afectividad política y democracia. *Revista de Estudios Políticos*, *173*, 27-54. https://doi.org/10.18042/cepc/rep.173.01

Bail, C.A. (2021). *Breaking the social media prism: How to make our platforms less polarizing*. Princeton University Press. https://doi.org/10.1515/9780691216508

Banaji, S. (2013). Everyday racism and «My tram experience»: Emotion, civic performance and learning on YouTube. [El racismo cotidiano y «Mi experiencia en un tranvía»: emoción, comportamiento cívico y aprendizaje en YouTube]. *Comunicar*, *40*, 69-78. https://doi.org/10.3916/C40-2013-02-07

Barberá, P. (2020). Social media, echo chambers, and political polarization. In N. Persily, & J. Tucker (Eds.), *Social media and democracy: The state of the field, prospects for reform* (pp. 34-55). Cambridge University Press. https://doi.org/10.1017/9781108890960

Berners-Lee, T. (2000). *Tejiendo la red*. Siglo XXI de España. https://bit.ly/3wZ1NMx

Berrocal-Gonzalo, S., Campos-Domínguez, E., & Redondo-García, M. (2014). Media prosumers in political communication: Politainment on YouTube. [Prosumidores mediáticos en la comunicación política: El «politainment» en YouTube]. *Comunicar*, *43*, 65-72. https://doi.org/10.3916/C43-2014-06

Bishop, S. (2018). Anxiety, panic and self-optimization: Inequalities and the YouTube algorithm. *Convergence*, *24*, 69-84. https://doi.org/10.1177/1354856517736978

Castells, M. (2001). *La era de la información: Economía, sociedad y cultura*. Alianza Editorial. https://bit.ly/3LXI18w

Chadwick, A. (2009). Web 2.0: New challenges for the study of e-democracy in an era of informational exuberance. I/S: A. *Journal of Law and Policy for the Information Society*, *5*(1), 9-41. https://bit.ly/3MZopSH

Chen, A., Nyhan, B., Reifler, J., Robertson, R., & Wilson, C. (2021). *Exposure to alternative & extremist content on YouTube*. Anti-Defamation League. https://bit.ly/3MZ19E9

Covington, P., Adams, J., & Sargin, E. (2016). Deep neural networks for YouTube recommendations. In S. Sen, & W. Geyer (Eds.), *Proceedings of the 10th ACM Conference on Recommender Systems* (pp. 191-198). Association for Computing Machinery. https://doi.org/10.1145/2959100.2959190

Davidson, J., Livingston, B., Sampath, D., Liebald, B., Liu, J., Nandy, P., Van-Vleet, T., Gargi, U., Gupta, S., He, Y., & Lambert, M. (2010). The YouTube video recommendation system. In X. Amatriain, M. Torrens, P. Resnick, & M. Zanker (Eds.), *Proceedings of the fourth ACM conference on Recommender Systems* (pp. 293-296). Association for Computing Machinery. https://doi.org/10.1145/1864708.1864770

Demsar, J., Curk, T., Erjavec, A., Gorup, C., Hocevar, T., Milutinovic, M., Mozina, M., Polajnar, M., Toplak, M., Staric, A., Stajdohar, M., Umek, L., Zagar, L., Zbontar, J., Zitnik, M., & Zupan, B. (2013). Orange: Data mining toolbox. *Python*. *The Journal of Machine Learning Research*, *14*(1), 2349-2353. https://bit.ly/3pMIPBR

Dimopoulos, G., Barlet-Ros, P., & Sanjuas-Cuxart, J. (2013). Analysis of YouTube user experience from passive measurements. In *Proceedings of the 9th International Conference on Network and Service Management (CNSM 2013)* (pp. 260-267). IEEE. https://doi.org/10.1109/CNSM.2013.6727845

Goodrow, C. (2021). *On YouTube's recommendation system*. Blog YouTube. https://bit.ly/3wWAxhA

Habermas, J. (1981). *Historia y crítica de la opinión pública*. Gustavo Gili. https://bit.ly/3O0JOv1

Hernández, E., Anduiza, E., & Rico, G. (2021). Affective polarization and the salience of elections. *Electoral Studies*, *69*, 102203. https://doi.org/10.1016/j.electstud.2020.102203

Howard, J.W. (2021). Extreme speech, democratic deliberation, and social media. In C. Véliz (Ed.), *The Oxford Handbook of Digital Ethics* (pp. 1-22). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198857815.013.10

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S.J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, *22*, 129-146. https://doi.org/10.1146/annurev-polisci-051117-073034

Latorre, M. (2022). *Historia de la Web, 1.0, 2.0, 3.0 y 4.0*. Blog Marino Latorre. https://bit.ly/38un7QH

Lilleker, D.G., & Jackson, N. (2008). *Politicians and Web 2.0: The current bandwagon or changing the mindset?* [Conference]. Politics: Web 2.0 International Conference.

Luengo, O., García-Marín, J., & Blasio, E. (2021). COVID-19 on YouTube: Debates and polarisation in the digital sphere. [COVID-19 en YouTube: Debates y polarización en la esfera digital]. *Comunicar*, *69*, 9-19. https://doi.org/10.3916/C69-2021-01

Mcluhan, H.M. (1959). Myth and mass media. *Daedalus*, *88*(2), 339-348. https://bit.ly/3GtIs9v

Messina, J.P. (2022). *New directions in the ethics and politics of speech*. Routledge. https://doi.org/10.4324/9781003240785

Mohan, N. (2022). *Inside responsibility: What's next on our misinfo efforts*. Blog YouTube. https://bit.ly/38XAngS

Nielsen, R., & Fletcher, R. (2020). Democratic creative destruction? The Effect of a changing media landscape on democracy. In N. Persily, & J. Tucker (Eds.), *Social media and democracy: The state of the field, prospects for reform* (pp. 139-162). Cambridge University Press. https://doi.org/10.1017/9781108890960.008

O'Reilly, T., & Battelle, J. (2009). *Web squared: Web 2.0 five years on*. O'Reilly Media. https://bit.ly/3wYLBuG

Pariser, E. (2017). *El filtro burbuja: Cómo la web decide lo que leemos y lo que pensamos*. Taurus. https://bit.ly/3x0UyDX

Rasmussen, S.H.R., & Petersen, M. (2022). *From echo chambers to resonance chambers: How offline political events enter and are amplified in online networks*. PsyArXiv. https://doi.org/10.31234/osf.io/vzu4q

Rekoff, M.G. (1985). On reverse engineering. *IEEE Transactions on Systems, Man, and Cybernetics*, *15*(2), 244-252. https://doi.org/10.1109/TSMC.1985.6313354

Serrano-Contreras, I., García-Marín, J., & Luengo, O.G. (2020). Measuring online political dialogue: Does polarization trigger more deliberation? *Media and Communication*, *8*, 63-72. https://doi.org/10.17645/mac.v8i4.3149

Sunstein, C.R. (2007). *Republic.com 2.0*. Princeton University Press. https://bit.ly/3a3YFG8

Terren, L., & Borge-Bravo, R. (2021). Echo chambers on social media: A systematic review of the literature. *Review of Communication Research*, *9*, 99-118. https://doi.org/10.12840/ISSN.2255-4165.028

Tufekci, Z. (2018). YouTube, the great radicalizer. *The New York Times*. https://nyti.ms/38VTs2Y

Van-Bavel, J.J., Rathje, S., Harris, E., Robertson, C., & Sternisko, A. (2021). How social media shapes polarization. *Trends in Cognitive Sciences*, *25*(11), 913-916. https://doi.org/10.1016/j.tics.2021.07.013

Wigand, R., Wood, J., & Mande, D. (2010). *Taming the social network jungle: From Web 2.0 to social media*. [Conference]. AMCIS 2010 Proceedings. https://bit.ly/3NJF3WI

Yesilada, M., & Lewandowsky, S. (2022). Systematic review: YouTube recommendations and problematic content. *Internet Policy Review*, (1), 11-11. https://doi.org/10.31234/osf.io/6pv5c

Comunicar 74

Kaleidoscope

Researchs

Studies

Proposals

AlfaMed

EuroAmerican Interuniversity Research Network
on Media Literacy for Citizenship

www.redalfamed.org

© ISSN: 1134-3478 • e-ISSN: 1988-3293

Comunicar, n. 74, v. XXXI, 2023 | Media Education Research Journal | ISSN: 1134-3478; e-ISSN: 1988-3478
www.comunicarjournal.com

News literacy and online news between Egyptian and Spanish youth: Fake news, hate speech and trust in the media

Alfabetización periodística entre jóvenes egipcios y españoles: Noticias falsas, discurso de odio y confianza en medios

Dr. Sally Samy-Tayie. Assistant Professor, Department of Media, The Arab Academy for Science, Technology and Maritime Transport (Egypt) (sallytayie@gmail.com) (https://orcid.org/0000-0003-0094-0863)
Dr. Santiago Tejedor. Professor, Department of Journalism and Communication Sciences, The Autonomous University of Barcelona (Spain) (santiago.tejedor@uab.cat) (https://orcid.org/0000-0002-5539-9800)
Dr. Cristina Pulido. Professor, Department of Journalism and Communication Sciences, The Autonomous University of Barcelona (Spain) (cristina.pulido@uab.cat) (https://orcid.org/0000-0001-8630-7529)

ABSTRACT

Exposure to harmful content, fake news, and hate speech, calls to question whether citizens are 'responsible' when it comes to their online news behavior. This exploratory study aims to positioning news literacy as a mediating factor toward fostering civic responsibility by investigating its role in enhancing youth's online news experience and preparing them against harmful content. News literacy is defined as a multi-structural construct with three main constituents: motivation, knowledge, and skills. The study's main goal is accomplished through delving into youth's online news experience, and assessing their level of news literacy through a mixed methods approach relying on focus group discussions as a qualitative method and survey as a quantitative method. The study is applied to a sample of youth from Egypt and a sample of youth from Spain, on whom comparative analysis is conducted. Findings reveal a positive correlation between news literacy and: the ability to identify fake news, the ability to identify hate speech, engagement with news, caring about the veracity of content before sharing, and motivation to seek news. Negative perceptions about professional news media's performance have been expressed by respondents from both countries, with the main justification being that news media are not fulfilling their role as should be.

RESUMEN

La exposición a contenidos nocivos, noticias falsas y discursos de odio hace que nos preguntemos si los ciudadanos son «responsables» cuando hablamos de su comportamiento online. Este estudio exploratorio pretende posicionar la alfabetización periodística como un factor mediador hacia el fomento de la responsabilidad cívica, investigando su papel en la mejora de la experiencia de los jóvenes con las noticias online, preparándolos contra los contenidos nocivos. La alfabetización periodística se define como un concepto multiestructural con tres componentes principales: motivación, conocimientos y habilidades. Un objetivo del estudio es profundizar en la experiencia de los jóvenes con las noticias online, evaluando su nivel de alfabetización periodística mediante un enfoque de métodos mixtos que se basan en discusiones de grupos focales y en una encuesta. El estudio se aplica a una muestra de jóvenes de Egipto y a otra de jóvenes de España, a partir de las cuales se realiza un análisis comparativo. Los resultados revelan una correlación positiva entre la alfabetización periodística y: la capacidad de identificar las noticias falsas y el discurso de odio, el compromiso con las noticias, la preocupación por la veracidad del contenido antes de compartirlo y la motivación para buscar noticias. Los encuestados de ambos países han expresado percepciones negativas sobre la eficacia de los medios de comunicación profesionales, con la justificación principal de que los medios de comunicación no están cumpliendo su función como deberían.

KEYWORDS | PALABRAS CLAVE

News literacy, news, disinformation, social media, youth, media consumption.
fabetización periodística, noticias, desinformación, redes sociales, juventud, consumo de medios.

Received: 2022-05-29 | Reviewed: 2022-06-15 | Accepted: 2022-07-13 | OnlineFirst: 2022-10-30 | Published: 2023-01-01
DOI https://doi.org/10.3916/C74-2023-06 | Pages: 69-81

1. Introduction

Although disinformation and misinformation existed historically in the media environment, the current state of "information disorder", as named by Wardle and Derakhshan (2018), is unprecedented. It is having an impact on an individual basis, and on societies and countries, symptomized by division and polarization. March 2020 marked the eruption of a global crisis, which began as a health emergency. Then, the World Health Organization as well as UNESCO announced that we are simultaneously witnessing a "massive infodemic" referring to the waves of COVID-19-related disinformation (Posetti & Bontcheva, 2020). "Vaccine hesitancy" has been one of the major consequences of disinformation, and conspiracy theories, and was reported by 90% of the countries worldwide (Schiavo, 2020). With the chaotic online information environment charged with fake news and hate speech, created with the intention to misleadingly manipulate public opinion, news literacy is positioned as a weapon against harmful content online and as a key toward a responsible news experience through which youth can make informed decisions and become more politically engaged. Scholars support the importance of critical analysis and evaluation of online material to face the impact of harmful content; fake news (Pérez-Tornero et al., 2018; Lee, 2018; McGrew et al., 2017; Lotero-Echeverri et al., 2018; Rosenzweig, 2017) and/or hate speech (Gagliardone et al., 2015; Livingstone et al., 2008; Daniels, 2008).

An experiment conducted by Adjin-Tettey (2022) shows that those who received Media and Information Literacy training show a higher ability to detect disinformation and are less likely to share content impulsively. Accordingly, with the current news environment loaded with such harmful content, the study intends to understand how news literacy relates to this context. Are news literate individuals more responsible news users? Can news literacy become a weapon against harmful content online? Stemming from these ideas, conducting this study on samples from two completely different countries, with different cultures, languages, and democratic status; Egypt and Spain, provides essential implications for designing news and media literacy interventions for youth empowerment. The two countries are to a great extent, dissimilar with respect to geographical location and hence culture. Egypt is an African and Arab country with a population that exceeds 105 million. Spain is a European country with a population that exceeds 40 million. This exploratory study is developed with the main goal of positioning news literacy as a mediating factor toward fostering civic responsibility, by investigating its role in enhancing youth's online news experience, and preparing them against harmful content intertwined with the current online information environment. To investigate that, the study addresses the following research questions:

- RQ1: What is the level of news literacy of Spanish and Egyptian youth?
- RQ2: What are the patterns of news use on social media among Spanish and Egyptian youth?
- RQ3: How far are youth exposed to harmful content online and how do they perceive it?
- RQ4: What is the level of news engagement of youth from both countries?
- RQ5: How far do youth from Egypt and Spain trust professional news media?

2. Methodology

The study relies on a mixed-method approach, using the survey as a quantitative method and the focus group discussions as a qualitative one. Results from the qualitative study directed the design of the questionnaire.

The questionnaire and focus group discussion session plan were validated by consulting academic experts. The universe of the study is defined as university students who are social media users coming from media and communication academic backgrounds. Accordingly, these criteria were considered for selecting a purposive sample, which includes Egyptian and Spanish youth in the age category of 16 to 25.

Regarding the survey, being an exploratory study, the questionnaire was distributed to a sample of 110 students from Cairo University, Egypt and a sample of 74 students from the Autonomous University of Barcelona (UAB), Spain, both being highly ranked universities in big cities. The demographics of respondents from Egypt are as follows: Gender; 85% are females and 15% are males, Age; 98% of respondents are in the age group of more than 20 to 25, with 2% in the age group of 16 to 20. As for the demographics of respondents from Spain: Gender; 66% are females and 34% are males, Age; 57% are in the age group of more than 20 to 25, with 43% in the age group of 16 to 20.

Regarding the focus group discussions, two pilot focus group discussions were conducted with students from two public universities in Egypt and one more conducted with students from a private university: The Arab Academy for Sciences, Technology and Maritime Transport. One focus group discussion was conducted with students from UAB, Spain. Respondents in the focus group discussions were ten in each; half males and half females. Being an exploratory investigation, the study used a purposive sample with a number of participants less than that required for the results to be representative. It is also important to mention that given some access difficulties, the number of participants from Spain is fewer than that of participants from Egypt.

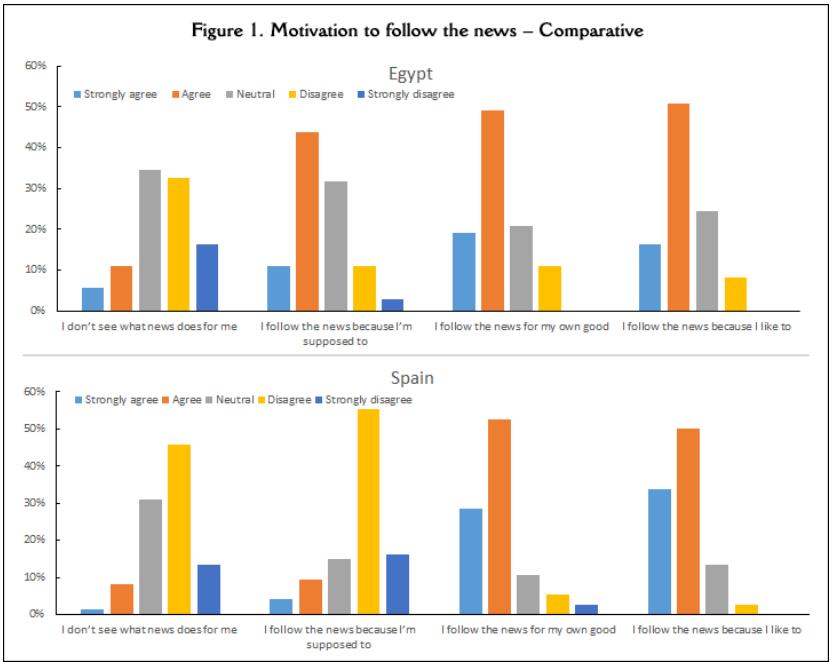
3. Findings: Youth's online news experience

Highlighting the similarities and differences between both countries, this section is divided into sections of pertinence to the main research question.

3.1. High levels of news literacy

News Literacy, regarded as a multi-dimensional construct, was measured by measuring motivation and knowledge areas based on Maksl et al. (2015) News Literacy Scale which is based on Potter's (2004) Cognitive Theory of Media Literacy, in addition to measuring skills of credibility assessment and verification based on Flanagin and Metzger (2000).

Figure 1. Motivation to follow the news – Comparative



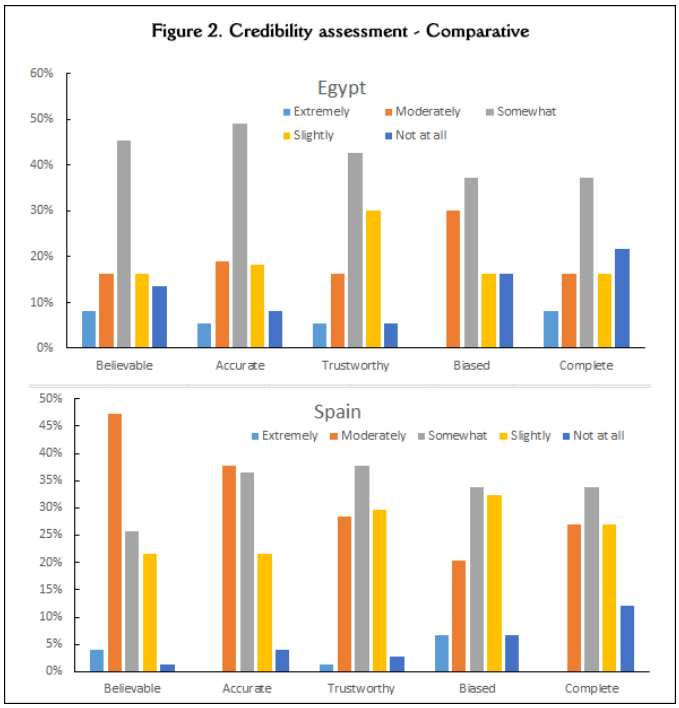
Based on the findings demonstrated, most respondents in both countries showed high levels of news literacy, assessed in terms of motivation, knowledge and skills. Based on the Likert scale statements used to assess respondents' motivation, youth from both countries demonstrate high levels of motivation to seek news. More than half of the Egyptian respondents (50.9%) agreed that they follow the news because they like to, with 16.4% strongly agreeing to the statement. Also, almost half of the Egyptian respondents (49.1%) agreed to the statement that they follow the news for their own good with 19.1% strongly agreeing. Half of the Spanish respondents (50%) agreed that they follow the news because they like to¸ with more than a third (33.8%) strongly agreeing to the statement. More than half of the respondents (52.7%) agree to the statement that they follow the news for their own good, with almost a third (28.4%) strongly agreeing to the statement. Hence, youth from both countries are considerably motivated to follow the news. Regarding, knowledge areas, despite the general level being similar, differences could be noted when it comes to different knowledge areas.

- Knowledge of news content: using a Likert scale, more than a third of Egyptian respondents (37.3%) strongly agreed that news companies choose stories based on what will attract the biggest audience, with more than a third (32.7%) agreeing. The results displayed show that most respondents from both countries are aware about news content selection from both sides, the journalists' and the audience's.
- Knowledge of the news media industry: In this area of knowledge, similarities and differences could be found. Respondents from both countries show awareness of ownership's influence. Almost half of the Egyptian respondents (46.4%) agreed to the statement that the owner of a media company influences the content that is produced, with almost a third (27.3%) strongly agreeing. More than half (54.1%) of the Spanish respondents strongly agreed that the owner of a media company influences the content that is produced, with more than a third (33.8%) agreeing. Egyptian respondents demonstrate confusion between the roles, definition and impact of the different persons involved in the industry (reporters, producers, cameramen, anchors, etc.), unlike Spanish respondents who demonstrate awareness about the same area.
- Testing awareness about the concept of objectivity, more than a third of Egyptian respondents selected the correct answer to the question "One common criticism of the news is that it is not objective. What do people who make that criticism typically mean by it?"; 35.5% selected the meaning that the reporter puts his/her opinion in the story, with 26.4% stating that they don´t know. More than half (58.1%) of the Spanish respondents selected the correct answer to the same question, with 16.2% stating they don't know. Here, another difference is spotted, with Spanish respondents demonstrating stronger awareness about the meaning of objectivity and its impact on content. The implications of these differences are discussed as correlating with the level of trust in the media in the specified section below.
- Knowledge about the news media effects: Findings from Egypt and Spain indicate awareness of respondents that a considerable part of the process depends on the audience´s interpretation and not just on the meaning intended by the reporter/journalist. The majority of respondents support the statement that two people might see the same news stories and get different information from it, with 42.7% agreeing and 35.5% strongly agreeing
- Knowledge about the real world: Respondents from both countries demonstrated awareness about the constructed nature of news. For instance, the majority of respondents support the statement that news makes things more dramatic than they really are, with 59% agreeing and 16.4% strongly disagreeing.
- Knowledge of the self: The results from both countries show that youth believe that they control media influences but not entirely; they still perceive some influences as not fully under their control. Respondents believe they are in control of how far they are informed and knowledgeable about the world, and how they can possibly avoid being misinformed.

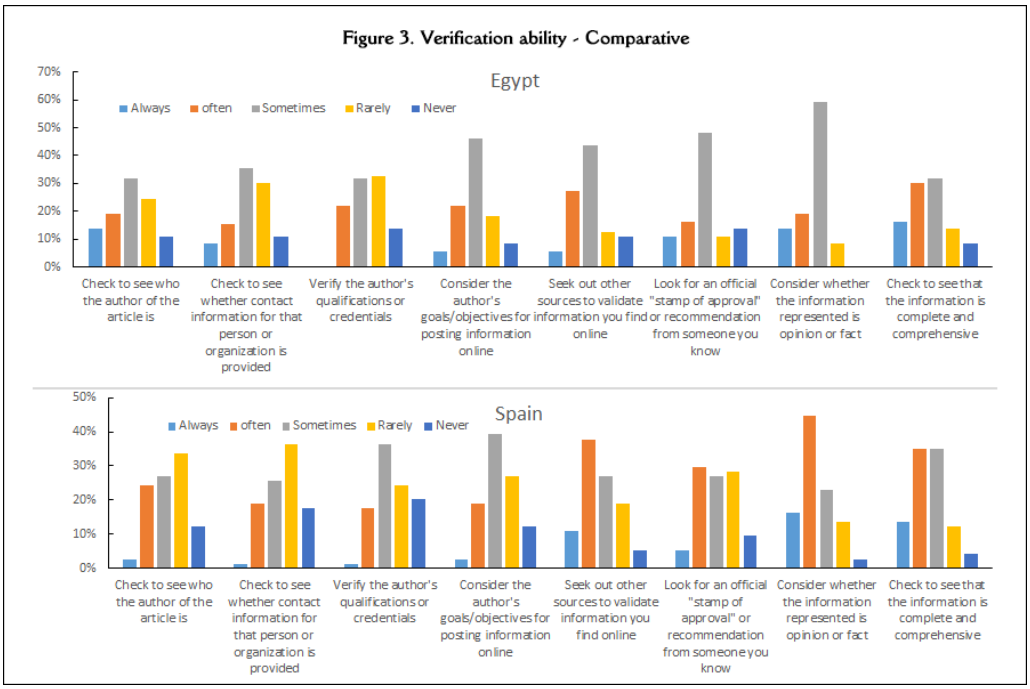
Regarding skills of credibility assessment and verification ability, there were differences between both countries. For credibility assessment, respondents were asked to assess the extent to which they find online news media credible by asking them to evaluate believability, accuracy, trustworthiness, bias, and completeness. Common among Egyptian respondents is the fact that the evaluations have almost all been average, with the biggest percentages of respondents selecting the response "somewhat" on the scale provided (with alternatives ranging from extremely to not at all).

Most responses in the case of Spain are dispersed among "moderately", "somewhat" and "slightly", except for believability where a more significant percentage of respondents choose "moderately" believable, as demonstrated. These findings go in line with those of media skepticism. Most Spanish respondents give responses more inclined towards a higher credibility assessment of the information provided by news media online. This especially applies when it comes to believability.

Taking into account that news media's websites are important sources for the sampled Spanish youth, the results are compatible. Regarding Egyptian respondents, doubtful stances were most prevalent going in accordance with their media trust findings.



In order to measure verification ability, respondents were asked about the frequency of doing specific actions when they read news online, by giving them statements to which they have to respond using a 5-point scale ranging from always to never.



Findings in the case of Egypt go in line with Flanagin and Metzger's (2000), with most Egyptian respondents verifying information online mostly sometimes or rarely. However, in the case of Spain, based on the findings, the verification activities need to be divided into content-related verification activities

and author-related ones. When it comes to author-related activities (such as checking the author's goals, verifying qualifications, etc.), Spanish respondents verify sometimes or rarely. On the other hand, when it comes to content-related verification activities, the majority of Spanish respondents' responses range from often to sometimes (such as seeking other sources, checking for completeness, verifying if the information is opinion or fact, etc.).

Based on the findings demonstrated, in both cases of Egypt and Spain, the majority of respondents demonstrate a high level of news literacy (65% in Egypt, 66% in Spain). Most respondents in both cases displayed motivation to seek news by supporting the statements that imply their understanding of the importance of news to them and rejecting statements that marginalize this role. Regarding knowledge areas, respondents from both countries demonstrate a good level of awareness about the concepts covered under each knowledge area. This applies to all, except for knowledge about the news industry; Egyptian youth demonstrated a weaker level than their Spanish counterparts in this area, specifically regarding the roles played by the different individuals involved in the news-making process. Differences between both countries exist when it comes to the tested skills. In the case of Egypt, findings for both (credibility assessment and verification ability) confirm the need to work on respondents' awareness about the essentiality to carry on certain inspections when reading news online, to be able to evaluate the credibility and verify the used information. In the case of Spain, the respondents appear to have a good level in both skills with the exception of author-related verification as highlighted.

Despite the close results between respondents from both countries, it is essential to note that *neutral* responses are more prevalent among Egyptian respondents than their Spanish counterparts. Scholars suggest that there are two possible scenarios behind reporting neutral stances; "true neutral"/"indifferent" or "don't know"/"undecided" (Raaijmakers et al., 2000). Accordingly, based on the nature of statements, it is suggested that neutral stances reported by Egyptian respondents can be analyzed as a lack of knowledge in support of Sturgis et al. (2012) findings; most respondents giving neutral answers were found to be either with no opinion or do not know. This has been demonstrated most in knowledge about news media effects results. On the other hand, Spanish respondents demonstrate better results when it comes to decisiveness by mostly reporting positive or negative stances.

3.2. Patterns of online news use and news engagement

Respondents from both countries displayed a preference for online sources of news over traditional media; there is especially a decline in newspaper readership. However, Spanish respondents' news use reveals a higher level of engagement with news than their Egyptian counterparts.

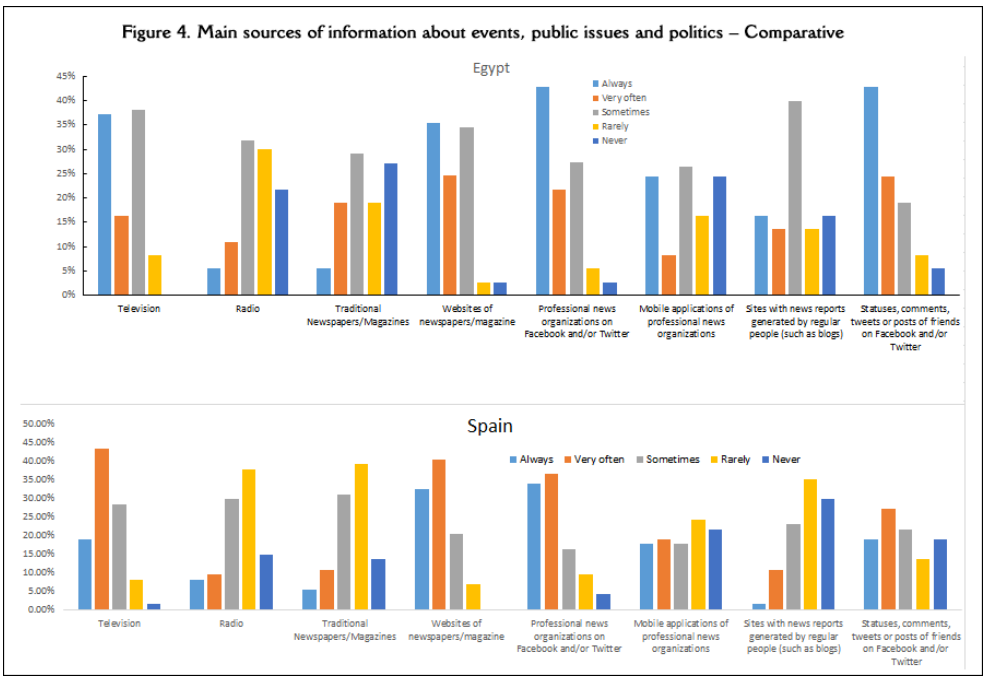
Egypt's findings show that the primary source(s) of news for respondents are accounts of professional news organizations on social media, followed by content generated by their peers on social media (statuses, tweets, comments, etc.) and websites of newspapers/magazines coming third. On the other hand, for Spanish respondents, newspaper/magazine websites come in the first place as a source of news, followed by accounts of professional news organizations on social media, then television comes third.

The fact that Egyptian respondents ranked professional news media's websites third as a source of news, leaves social media as the main news carrier. This implicates the ramifications that could result from relevant issues such as personalization algorithms based on which such platforms work. In other words, are social media/online news users aware of how content appears to them? According to empirical findings of a study done on university students in the US, youth "are largely unaware of whether and how news sources track user data and apply editorial judgments to deliver personalized results" (Powers, 2014).

In contrast with the findings from the Egyptian case, social media use for news does not come at the expense of professional news media's websites for Spanish respondents. This is reflected in the fact that such websites were ranked as the primary source of news for the sampled Spanish young people. This choice was followed by accounts of professional news organizations on social media, with peer-generated content pushed forth as a source of news. These findings reveal the extent to which Spanish youth are able to differentiate between social media as carriers of news content rather than sources, and professional sources to get their information. This supports Braun and Gillespie's (2011) statement about the importance that users realize such difference.

Analyzing youth's preference to get their news on social media, Hermida et al. (2012) state that: "the traditional gatekeeping function of the media is weakened as a significant proportion of news consumers turn to family, friends, and acquaintances to alert them to items of interest". This statement is supported by the qualitative findings of the study in the case of Egypt, where Egyptian respondents emphasized the importance of the opinions/views of trusted individuals. It further applies –in a weaker sense– in the case of Spanish respondents. A respondent from Egypt mentioned that she has become "less interested in getting …news from [the] TV, they lie and if not, they try to manipulate us... this is why checking opinions around us became much more important than before".

Figure 4. Main sources of information about events, public issues and politics – Comparative



When it comes to social media platforms, priorities are completely different in both countries. With Facebook and YouTube being the most frequently used among Egyptian respondents, Instagram and Twitter are most used by Spanish respondents. Since the study focuses on Facebook and Twitter, it is then essential to note that Facebook is massively declining among Spanish youth, and Twitter is minimally used among their Egyptian counterparts, based on the qualitative and quantitative findings for both countries. This is confirmed when asked about different news sources. Egyptian respondents reported Facebook as a major source of news, followed by Google News, and then YouTube; while Spanish respondents reported Twitter as a major source of news, followed by El Pais Online (newspaper), and then El Diario (newspaper).

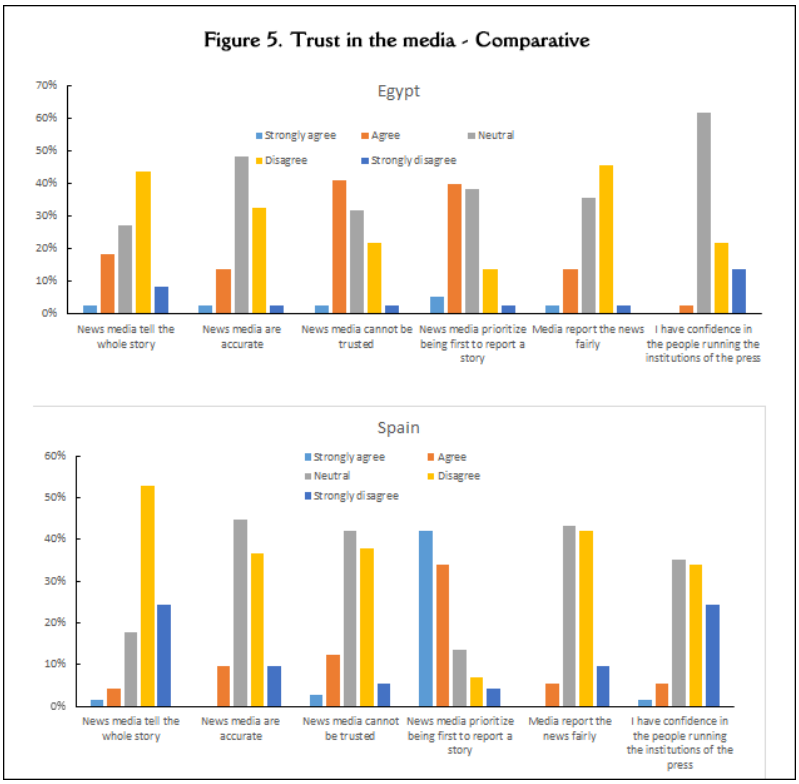
Scholars suggest that the decision to use one platform and not the other is related to major personality traits (Hughes et al., 2012). Accordingly, with a focus on information seeking, Facebook's users are believed to get their information while/through "socializing", while Twitter's users are believed to intentionally seek information for its utility and "value" (Hughes et al., 2012). Contradictorily, Egyptian respondents report Facebook's primary information use is to get news about current events from mainstream media; Spanish respondents barely agree. Moreover, Spanish respondents report Twitter's primary information use is getting news about current events from friends, similar to findings from Egypt. These findings contradict Hughes et al. (2012) aforementioned findings. However, it is important to note that in the focus group discussions, Egyptian respondents emphasized the importance of the views of others which justifies their dependence on social media (especially Facebook) for news. Spanish respondents also mentioned during the discussions that they check what their peers have to say about the different issues raised, which is well-matched with their first informational use of Twitter. Accordingly, Egyptian and

Spanish youth in the study demonstrated different levels of engagement with professional news media as shown.

3.3. Perception of news media, knowledge, and trust

Negative perceptions about professional news media's performance have been expressed by respondents from both countries, with the main justification being that news media are not fulfilling their role as should be. Respondents from both countries are well aware that the news media's main role is to inform the citizenry. As previously discussed, youth from both countries generally show a good level of knowledge in the five knowledge areas investigated. Findings from both countries support Arendt et al.'s (2016) media-related selection which highlights journalists' gatekeeping function as filters of content and the audience's selective exposure and sharing behavior. Media-related selection (MRS) is an attempt by Arendt et al. (2016) to combine three theoretical tenets of selection taking into account the two main actors in the news experience; journalists and users. The three theoretical concepts addressed under MRS are gatekeeping (journalists' selection), selective exposure (users' selection where confirmation bias is taken into account), and news sharing on social networking sites (journalists' and users' sharing behavior is considered). As previously demonstrated, three knowledge areas reflect MRS; knowledge of content, knowledge about media effects and knowledge about the world. Hence, in the three areas, respondents from both countries demonstrate similar results, being aware of how journalists' and audience's selection affects the product (the news story). Respondents showed awareness about factors that impact audience selection, such as cognitive dissonance (Festinger, 2009) and confirmation bias (Jonas et al., 2001) in qualitative and quantitative findings.

Figure 5. Trust in the media - Comparative



On the other hand, when it comes to knowledge of the news media industry, differences between both groups of respondents from Egypt and Spain emerged. Egyptian respondents demonstrate confusion between the roles, definitions, and impact of the different pople involved in the industry (reporters, producers, cameramen, anchors, etc.), unlike Spanish respondents who demonstrate awareness about the same area. Scholars (Craft et al., 2017; Pérez-Rodríguez & Delgado-Ponce, 2012) emphasize

the essentiality of knowledge about the media industry for trust and engagement with the news after establishing correlations between these variables. This study supports such correlations as shown in the findings on media skepticism.

By assessing respondents' trust in the media, it was found that Spanish respondents, despite being skeptical, show a better level of trust than their Egyptian counterparts. Egyptian respondents in the qualitative and quantitative responses demonstrate the essentiality of views of peers or trusted persons in their circles for them to be able to form an attitude or an opinion toward different issues. This comes at the expense of intentionally seeking information through official news websites; which further justifies depending on social media for news. On the other hand, Spanish respondents demonstrate a lack of trust in a more critical manner; consciousness about the reasons why they do not trust the media, yet they acknowledge that professional news media is what they should follow to remain up to date and informed. One of the Spanish respondents commented: "we should be aware that we only receive part of the occurrences that are happening as a result of what the journalists select ... I think that news contributes to the image that we have about the world; which gives us an image/a vision that doesn't perfectly reflect reality or the society". This is emphasized through their primary source of information being websites of newspapers/magazines, while still using social media and their friends' networks for elaboration and context.

In light of these findings, it is worth noting that the results from the different variables tested in each of the two cases are compatible with each other as displayed. Adding to that, reference to the argument on criticism versus cynicism is essential for the different consequences each has on news engagement and also political engagement. Being cynical refers to having a general judgment, usually negative in case of cynicism towards the media, based on which an individual takes a decision of disengagement. For example, the judgment that news media lie or present negative content all the time, and so I am not interested in following news anymore (Buckingham, 2000; Mihailidis, 2008).

3.4. Fake news and hate speech: Exposure and detection

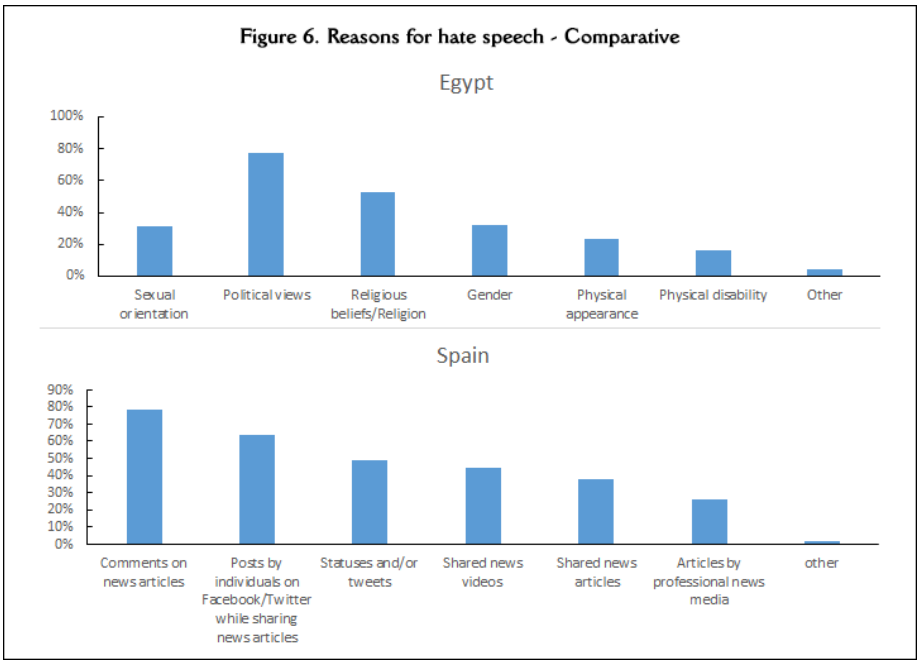
When asked whether or not they are aware that some news stories shared on social media are fake news, a massive majority of respondents from both countries (97.3% Egypt, 97.3% Spain) reported that they are aware of it. Also, the majority of the respondents support the statements that fake news on social media can cause confusion; it is important to check the credibility of a news article before sharing it.

Respondents displayed rejection of the statement; if the headline is interesting, I share without reading; sharing news stories shared by friends without reading; that they do not necessarily check the sources of information in a news story before they share it. When asked whether they believe they have the competencies to tell whether or not a news story is fake; Egyptian respondents' responses were more dispersed that their Spanish counterparts, who mostly supported the statement. In the case of Egyptian respondents, a discrepancy between their awareness of the problem and their actions could be noted, besides the prevalence of neutral responses indicating doubts. The percentages displayed reveal awareness from the side of the Egyptian youth about the potential impact of fake news and their reported perception that it is important to consider the source which delivers the news, which was barely reflected in the previously mentioned findings on verification ability. However, there is also a high percentage who reported neutral positions, in that sense, indicating confusion, indifference, or lack of enough awareness (like the case with sharing news stories only from professional sources, where more than a third stated being neutral about it).

In the case of Spanish respondents, in line with the findings on verification ability, the percentages displayed reveal that youth are more aware of the essentiality of double-checking content than of verifying sources. This has been reflected through their more dispersed responses to the statement about checking sources of information in a news story before sharing. In line with the qualitative findings, respondents show that they care about the credibility of information and understand their responsibility when sharing content.

Regarding their ability to detect fake news, Spanish respondents show more confidence in their competencies than their Egyptian counterparts with the majority reporting "neutral". Interestingly,

respondents from both countries reported that the hateful writings they encounter attack individuals/groups for political views most (77% Egypt, 86.5% Spain). Egyptian respondents encountering hate speech for religious beliefs/religion (52.7%) and gender (32.4%) come next. Spanish respondents encountering hate speech for gender (74.3%) and sexual orientation (70.3%) come after political views.



When asked about their reactions to harmful content (fake news or hate speech), the majority of respondents stated that they ignore it (50.9% in Egypt, 47.3% in Spain). Findings from both countries reveal that harmful content does shape a considerable part of youth's online news experience.

3.5. Reading, sharing and content generation

Scholars (Fletcher et al., 2018; Parlapiano & Lee, 2018) note that false material multiplies on social media much more than truthful content thanks to algorithms which makes individuals' exposure to content in line with their beliefs higher, creating a filter bubble (Allcott & Gentzkow, 2017; Dornan, 2017; Hawdon et al., 2015). Hence, many scholars agree that human behavior online complicates the problem of mis/dis-information (Vosoughi et al., 2018; Anderson & Rainie, 2017; Newman et al., 2017; Allcott & Gentzkow, 2017). Consequently, youth's news behavior was investigated by understanding how they read, share and generate content online.

Reading: According to the qualitative findings of both countries, Egyptian and Spanish respondents only read headlines of news stories on social media most of the time. Respondents from both countries demonstrate awareness that professional journalists in many instances use sensationalism and write misleading headlines (clickbait content) with the sole purpose of attracting users at the expense of content quality. Despite this, respondents still barely read full stories, unless it concerns "big events" as noted by Egyptian interviewees, or "interesting" to them as noted by their Spanish counterparts. It is paradoxical that, despite being aware and clear that fabricated news and misleading headlines are common on social media, still the majority of students just read the headlines and rarely clicks on them for full stories. Spanish respondents shed light on how reading nowadays has changed; "more like skimming through the overload of information encountered". Besides headlines, what else do young people read? Comments.

Based on the findings from both countries, minor percentages of respondents stated that they never read comments (8.2% in Egypt, 9.2% in Spain), which means that the majority do. Investigating their motivations (four categories), Egyptian and Spanish respondents were found to have different priorities, with seeking information being the primary motive for Egyptian respondents, it becomes clear how important the

role of peer citizens in creating content has become, and especially how it is regarded by a majority as complementary to the information provided by journalists through the news article itself. This should be taken into account bearing in mind how studies (Erjavec, 2014; Erjavec & Kova i , 2012) revealed that hate speech exists in many instances in comments on news, as previously highlighted. Furthermore, the problem is aggravated, with scholars emphasizing how hate speech producers use techniques of rewriting and reshaping the meanings in the news articles consistently to serve their purposes (Erjavec & Kova i , 2012). For Spanish respondents, social interaction is the primary motive to read comments. Hence, it becomes clear how peer citizens can impact each other's opinions by creating an online public sphere for deliberations. This means, in the context of this study, that being conscious of the creators of content is necessary not to be misled. Entertainment motivation comes second for Egyptian respondents, while information motivation is second for Spanish respondents.

Sharing: Motivations behind youth´s sharing behavior on social media being investigated. The motivations were divided into four categories: information seeking, socializing, entertainment and status-seeking (Lee & Ma, 2012). For respondents from both countries, information-seeking and socializing motivations are most influential. For information-seeking motivation, most respondents stated that sharing news on social media:

- Helps them store useful information (64.5% Egypt, 33.8% Spain).
- Helps them keep up to date on the latest news and events (32% Egypt, 50% Spain).
- It becomes easy to retrieve information when needed (41% Egypt, 33.8% Spain).

Socializing is also regarded as an essential motive for respondents to share information on social media; the sampled youth stated that:

- It is effective to exchange ideas with other people (44% Egypt, 39.2% Spain).
- It makes them keep in touch with people (40% Egypt, 13.5% Spain).
- It helps them interact with people when sharing news (30% Egypt, 47.3% Spain).

Entertainment and status-seeking were less important as motivations to share news in the cases of both countries. However, how likely are respondents to share in the first place? According to the quantitative findings, Egyptian respondents are more likely to share news than their Spanish counterparts as reflected by the percentage of participants reporting they never share news on social media (8.2% Egypt, 21.6% Spain). In contrast, based on the qualitative findings, most Egyptian interviewees demonstrated reluctance to share news on social media, mainly for fear of being judged for views and of unintentionally misinforming. While Spanish interviewees demonstrated the fear of unintentionally misinforming their community as the main reason for not sharing. Such findings go in line with the fact that youth from both countries demonstrated care about the veracity of any material they share.

4. Conclusion and discussion: Implications to consider

Overall, there are more similarities than differences between the findings of the Egyptian and Spanish cases. The same correlations were concluded in both cases as demonstrated in the findings and analysis. The most important differences between the findings of the two countries are in the level of engagement to news as concluded, based on differences in:

- Trust in the media; cynicism versus criticism.
- Knowledge about the news media industry.
- Primary sources of news.
- Informational use of social media platforms.

Furthermore, motivations to read were found to be different. Such a finding is related to primary sources of news; Egyptians' being social media and Spanish individuals' being professional media. This leads to a conclusion that Spanish respondents use social media for their original role as news disseminators/carriers and are aware of how socializing (and being informed while doing so) is its first purpose, rather than counting on it for obtaining information on which to build opinions.

When discussing hate speech, one common argument between respondents from both countries is the fact that media polarizes and separates people by using stereotypes, labels and the 'us and them' rhetoric. Such polarization exists in both Egyptian and Spanish societies.

In light of such findings, youth from both countries suffer an environment charged with misleading content which is a symptom that potentially affects citizens' political engagement negatively. According to scholars, the two main prerequisites of a healthy democracy are the dissemination of relevant facts and information by politicians and the media, and citizens' use of such information in a manner that suits their preferences and, at the same time, "correct mistaken conceptions" (Kuklinski et al., 2000). With the majority of youth in both countries demonstrating the existence of confirmation bias in their news/information-seeking behavior, the task of the news to "correct mistaken conceptions" is more challenging than ever before. This especially applies to levels of media trust playing an essential role; the lack of it leads to the lack of engagement with the news.

Now more than ever, quality journalism became a need for survival. In that sense, involving current and future journalists in the field of news literacy is essential to provide a practical perspective, regain users' trust and reflect on their own work. The first step is admitting that the current online environment has taken its toll on the quality of journalism. Journalists are under pressures such as the immediacy of breaking news versus accuracy and verification, and audience trust versus maintaining their economic survival. This inevitably aggravates the information chaos problem by allowing false and misleading content to grow. Respondents in this study accuse the professional media of inciting hatred primarily for political views and hence dividing the people and labeling them. Hence, news literacy is deemed essential for journalists to incorporate and regain their image as legitimate sources of information to citizens by proving transparency and educating the public about their work. Setting the theoretical grounds for news literacy is still underway. Hence further research is necessary in order to seek defining and examining theoretical tenets for news literacy.

Authors' Contribution

Idea, S.T; Literature Review (state of the art), S.T, C.P, S.T.; Methodology, S.T; Data analysis, S.T, S.T, C.P; Results, S.T.; Discussion and conclusions, S.T, S.T, C.P; Writing (original draft), S.T; Final revisions, S.T; Project design and funding agency, S.T.

Funding Agency

The research has been carried out through funding obtained as a Consolidated Research Group Funded by the Agency for University and Research Grants of the Generalitat de Catalunya with file number 2017 SGR 1504.

References

Adjin-Tettey, T.D. (2022). Combating fake news, disinformation, and misinformation: Experimental evidence for media literacy education. *Cogent Arts & Humanities*, *9*(1), 2037229. https://doi.org/10.1080/23311983.2022.2037229

Allcott, M., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*(2), 211-236. https://doi.org/10.1257/jep.31.2.211

Anderson, J., & Rainie, L. (2017). *The future of truth and misinformation online*. Pew Research Center. https://pewrsr.ch/3PPlk9e

Arendt, F., Steindl, N., & Kümpel, A. (2016). Implicit and explicit attitudes as predictors of gatekeeping, selective exposure, and news sharing: Testing a general model of media-related selection. *Journal of Communication*, *66*(5), 717-740. https://doi.org/10.1111/jcom.12256

Braun, J., & Gillespie, T. (2011). Hosting the public discourse, hosting the public: When online news and social media converge. *Journalism Practice*, *5*(4), 383-398. https://doi.org/10.1080/17512786.2011.557560

Buckingham, D. (2000). *The Making of Citizens: Young People, News and Politics*. Routledge. http://doi.org/10.4324/9780203132272

Craft, S., Ashley, S., & Maksl, A. (2017). News media literacy and conspiracy theory endorsement. *Communication and the Public*, *2*, 388-401. https://doi.org/10.1177/2057047317725539

Daniels, J. (2008). *Race, civil rights, and hate speech in the digital era*. MIT Press. https://bit.ly/3x1QX7m

Dornan, C. (2017). *Dezinformatsiya: The past, present and future of fake news. Series of reflection papers*. Canadian Commission for UNESCO. https://bit.ly/3m2CVx8

Erjavec, K. (2014). Readers of online news comments: Why do they read hate speech comments? Annales Histoire. *Sciences Sociales*, (3), 451-462. https://bit.ly/3x1Rh64

Erjavec, K., & Kova i , M.P. (2012). You don´t understand, this is a new war! Analysis of hate speech in news web sites' comments. *Mass Communication and Society*, *15*, 899-920. https://doi.org/10.1080/15205436.2011.619679

Festinger, L. (2009). Cognitive dissonance theory. In E. Griffin, & A. L. G. Sparks (Eds.), *A first look at communication theory* (pp. 194-206). McGraw-Hill. https://bit.ly/3zltnDF

Flanagin, A.J., & Metzger, M.J. (2000). Perceptions of internet information credibility. *Journalism & Mass Communication Quarterly*, *77*(3), 515-545. https://doi.org/10.1177/107769900007700304

Fletcher, R., Cornia, A., Graves, L., & Nielsen, R.K. (2018). *Measuring the reach of fake news and online distribution in Europe*. Reuters Institute, Unviersity of Oxford. https://bit.ly/3GDDtD3

Gagliardone, I., Gal, D., Alves, T., & Martinez, G. (2015). *Countering online hate speech. The United Nations Educational, Scientific and Cultural Organization (UNESCO)*. https://bit.ly/3GDDtD3

Hawdon, J., Oksanen, A., & Räsänen, P. (2015). Online extremism and online hate: Exposure among adolescents and young adults in four nations. *Nordicom-Information*, *37*, 29-37. https://bit.ly/3BAbzaN

Hermida, A., Fletcher, F., Korell, D., & Logan, D. (2012). Share, like, recommend: Decoding the social media news consumer. *Journalism Studies*, *13*(5), 815-824. https://doi.org/10.1080/1461670x.2012.664430

Hughes, D.J., Rowe, M., Batey, M., & Lee, A. (2012). A tale of two sites: Twitter vs. facebook and the personality predictors of social media usage. *Computers in Human Behavior*, *28*(2), 561-569. https://doi.org/10.1016/j.chb.2011.11.001

Jonas, E., Schultz-Hardt, S., Frey, D., & Thelen, N. (2001). Confirmation bias in sequential information search after preliminary decisions: An expansion of dissonance theoretical research on selective exposure to information. *Journal of Personality and Social Psychology*, *80*(4), 557-571. https://doi.org/10.1037/0022-3514.80.4.557

Kuklinski, J.H., Quirk, P.J., Jerit, J., Schwieder, D., & Rich, R.F. (2000). Misinformation and the currency of democratic citizenship. *The Journal of Politics*, *62*(3), 790-816. https://doi.org/10.1111/0022-3816.00033

Lee, C.S., & Ma, L. (2012). News sharing in social media: The effect of gratifications and prior experience. *Computers in Human Behavior*, *28*, 331-339. https://doi.org/10.1016/j.chb.2011.10.002

Lee, N.M. (2018). Fake News, phishing, and fraud: A call for research on digital media literacy education beyond the classroom. *Communication Education*, *67*(4), 460-466. https://doi.org/10.1080/03634523.2018.1503313

Livingstone, S., Couvering, V.E., & Thumin, N. (2008). Converging traditions of research on media and information literacies: Disciplinary, critical, and methodological issues. In J. Coiro, M. Knobel, C. Lankshear, & D. J. Leu (Eds.), *Handbook of Research on New Literacies* (pp. 103-132). Routledge. https://doi.org/10.4324/9781410618894-12

Lotero-Echeverri, G., Romero-Rodríguez, L.M., & Pérez-Rodríguez, M.A. (2018). Fact-checking vs. fake news: Confirmation journalism as a tool of media literacy against misinformation. *Index.Comunicación*, *8*(2), 295-316. https://bit.ly/3oIXxvC

Maksl, A., Ashley, S., & Craft, S. (2015). Measuring news media literacy. *Journal of Media Literacy Education*, *6*(3), 29-45. https://doi.org/10.23860/jmle-6-3-3

Mcgrew, S., Ortega, T., Breakstone, J., & Wineburg, S. (2017). The challenge that's bigger than fake news: Civic reasoning in a social media environment. *American Educator*, *41*(3), 4-9. https://bit.ly/3NNynXa

Mihailidis, P. (2008). *Beyond cynicism: How media literacy can make students more engaged citizens*. College Park: University of Maryland, College Park. https://bit.ly/3ze5Rbu

Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D.A.L., & Nielsen, R.K. (2017). *Reuters institute digital news report 2017*. Reuters Institute. https://bit.ly/3x0mq9R

Parlapiano, A., & Lee, J.C. (2018). The propaganda tools used by russians to influence the 2016 election. *The New York Times*. https://nyti.ms/3x4navR

Pérez-Rodríguez, M.A., & Delgado-Ponce, A. (2012). From digital and audiovisual competence to media competence: Dimensions and indicators. [De la competencia digital y audiovisual a la competencia mediática: Dimensiones e indicadores]. *Comunicar*, *39*, 25-34. https://doi.org/10.3916/C39-2012-02-02

Pérez-Tornero, J.M., Tayie, S.S., Tejedor, S., & Pulido, C. (2018). How to confront fake news through news literacy? State of the art. *Doxa.Comunicación*, *26*, 211-235. https://doi.org/10.31921/doxacom.n26a10

Posetti, J., & Bontcheva, K. (2020). *Disinfodemic: Deciphering COVID-19 disinformation*. UNESCO. https://bit.ly/3NPrSn

Potter, W.J. (2004). *Theory of media literacy: A cognitive approach*. Sage Publications. https://doi.org/10.4135/9781483328881

Powers, E.M. (2014). *How students access, filer and evaluate digital news: Choices that shape what they consume and the implications for news literacy education*. [Doctoral Disseration, University of Maryland]. https://bit.ly/3PXhEmq

Raaijmakers, Q.A., Van Hoof, J.T.C., Hart, H., Verbogt, T.F.M.A., & Vollebergh, W.A. (2000). Adolescents' midpoint responses on Likert-type scale items: Neutral or missing values? *International Journal of Public Opinion Research*, *12*, 208-216. https://doi.org/10.1093/ijpor/12.2.209

Rosenzweig, A. (2017). Understanding and undermining fake news from the classroom. *Berkeley Review of Education*, *7*(1), 105-117. https://doi.org/10.5070/b87110055

Schiavo, R. (2020). Vaccine communication in the age of COVID-19: Getting ready for an information war. *Journal of Communication in Healthcare*, *13*(2), 73-75. https://doi.org/10.1080/17538068.2020.1778959

Sturgis, P., Roberts, C., & Smith, P. (2012). Middle alternatives revisited: How the neither/nor response acts as a way of saying "I don't know. *Sociological Methods & Research*, *43*(1), 15-38. https://doi.org/10.1177/0049124112452527

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146-1151. https://doi.org/10.1126/science.aap9559

Wardle, C., & Derakhshan, H. (2018). Thinking about information disorder: Formats of misinformation, disinformation and mal-informatio. In C. Ireton, & J. Posetti (Eds.), *Journalism, fake news and disinformation: Handbook for journalism, education and training* (pp. 44-55). UNESCO. https://bit.ly/3wZOPwN

© ISSN: 1134-3478 • e-ISSN: 1988-3293

Comunicar, n. 74, v. XXXI, 2023 | Media Education Research Journal | ISSN: 1134-3478; e-ISSN: 1988-3478
www.comunicarjournal.com

Factors determining the use of e-learning and teaching satisfaction

Factores determinantes en el uso del e-learning y la satisfacción docente

Dr. Ana-Maria Cazan. Professor, Department of Psychology, Education and Teacher Training, Transilvania University of Brasov (Romania) (ana.cazan@unitbv.ro) (https://orcid.org/0000-0003-4521-702X)
Dr. Catalin-Ioan Maican. Professor, Department of Management and Economic Informatics, Transilvania University of Brasov (Romania) (maican@unitbv.ro) (https://orcid.org/0000-0002-1968-9958)

ABSTRACT

Even though in 2021 many universities have decided to resume teaching activities face to face, we believe that the use of online applications will remain a feature of the educational system due to the flexibility offered and the learning possibilities. We aim to analyze the predictive role of personal factors, such as self-efficacy, technostress creators, technostress inhibitors, and tolerance to uncertainty in the use of e-learning tools for teaching and in the use of these applications in the context of the uncertainty generated by the pandemic. The sample consisted of 1,517 academics. The results showed that technostress creators mediate the relationships between technostress inhibitors, technology self-efficacy, use of applications, and satisfaction with the use of e-learning platforms. Although the current context is dominated by uncertainty, the hypotheses regarding the direct and indirect effects of uncertainty in the use of online s in education were partially sustained. The most important finding of our study is that, although the current context is characterized by uncertainty, the negative impact of the resulting higher levels of stress can be counteracted by a high level of technology self-efficacy which, in turn, predicts a greater extent the use of platforms and the satisfaction of using these platforms.

RESUMEN

Si bien en 2021 muchas universidades han decidido retomar la actividad docente presencial, creemos que el uso de aplicaciones en línea seguirá siendo una característica del sistema educativo por la flexibilidad que ofrece y las posibilidades de aprendizaje. Nuestro objetivo es analizar el papel predictivo de factores personales, como la autoeficacia, los creadores de tecnoestrés, los inhibidores del tecnoestrés y la tolerancia a la incertidumbre sobre el uso de herramientas de e-learning para la enseñanza y sobre el uso de estas aplicaciones en el contexto de la incertidumbre generada por la pandemia. La muestra estuvo conformada por 1.517 académicos. Los resultados mostraron que los creadores de tecnoestrés median las relaciones entre inhibidores de tecnoestrés, autoeficacia tecnológica, uso de aplicaciones y satisfacción hacia el uso de plataformas de e-learning. Aunque el contexto actual está dominado por la incertidumbre, las hipótesis sobre los efectos directos e indirectos de la incertidumbre sobre el uso de la aplicación en línea en la educación se sustentaron parcialmente. El hallazgo más importante de nuestro estudio es que, aunque el contexto actual se caracteriza por la incertidumbre, el impacto negativo de los mayores niveles de estrés resultantes puede ser contrarrestado por un alto nivel de autoeficacia tecnológica que, a su vez, predice en mayor medida el uso de plataformas y la satisfacción de usar estas plataformas.

KEYWORDS | PALABRAS CLAVE

Technology self-efficacy, technostress creators, technostress inhibitors, intolerance to uncertainty, e-learning, satisfaction.
Autoeficacia tecnológica, creadores de tecnoestrés, inhibidores del tecnoestrés, intolerancia a la incertidumbre, e-learning, satisfacción.

Received: 2022-03-05 | Reviewed: 2022-06-27 | Accepted: 2022-07-11 | OnlineFirst: 2022-10-30 | Published: 2023-01-01
DOI https://doi.org/10.3916/C74-2023-07 | Pages: 83-93

1. Introduction and state of the art

The spread of the coronavirus epidemic generated uncertainty concerning academic life, with nearly 85% of the total enrolled learners in 172 countries being affected (UNESCO, 2020). By the fall of 2020, many universities identified solutions to combine or replace in-person instruction with online instruction, and by the winter of 2020-2021, most universities provided online-only classes based on ICT systems (UNESCO, 2020). Some of these systems offer blended learning, others are of the massive online open courses type, and others only offer a simple platform through which teachers can share content with students, offering video conferencing features to replace in-person classes. Moving all programs online proved challenging for all of them, and many institutions have had to quickly find solutions to replace in-class education. Recent research found that barriers limiting how technology can enhance teaching at universities still exist (Polly et al., 2021).

Although some teachers have positive attitudes towards digital tools, they still need training in the use of ICT tools in teaching and content creation (Casero-Béjar & Sánchez-Vera, 2022) since, even though in 2021 many universities have decided to resume face-to-face teaching activities, the use of online applications will remain a feature of the educational system due to the flexibility offered and the learning opportunities (Hodges et al., 2020). This flexibility is dependent, among other factors, on the characteristics of teachers, such as self-efficacy, technostress, and tolerance to uncertainty, including the context in which the teaching activities take place. In these conditions, the main research question is: are the personal factors affecting the use of e-learning tools for teaching in conditions of uncertainty generated by the pandemic? This research question was tested on a representative sample of Romanian academics from the most important higher education institutions. The following section will present determinants of using e-learning tools among teachers including self-efficacy, technostress creators, inhibitors, and uncertainty, sustained by relevant recent studies.

1.1. Determinants of e-learning satisfaction among university teachers

Mouakket and Bettayeb (2015) found that the perceived usefulness of e-learning is the most important predictor of teachers' satisfaction with e-learning. Satisfaction in online teaching is correlated with perceived effectiveness (Almuwais et al., 2021), even though teacher training has no influence on satisfaction (Al-Samarraie et al., 2018). Ease of use, accessibility of IT Infrastructure (Rokhimah & Sirait, 2021), job-related factors (Marasi et al., 2022), support for flexibility in teaching schedules and appropriate training are more likely to make teachers satisfied with online teaching (Stickney et al., 2019). While these aspects have been debated as factors determining e-learning satisfaction among university teachers, fewer studies have highlighted the role of factors such as technology self-efficacy, technostress, and intolerance to uncertainty. Nevertheless, let's not forget that an e-learning platform is just a tool whose results depend on the teachers' involvement, their teaching skills, or their abilities to adapt.

1.2. Techno self-efficacy (TSE)

According to Bandura's social-cognitive theory, self-efficacy is an important factor in predicting task performance (Bandura, 1997), influencing individuals' emotional reactions or thought patterns in stressful situations. Research has found that highly perceived technology self-efficacy encourages the use of computers while reducing an individual's IT-related anxiety (Pressley & Ha, 2021), generating a constructive atmosphere for supporting efforts and easier adaptation to IT related changes (Bakar et al., 2018). For teachers, TSE is related to the effective use of platform facilities. Research has tested several factors related to technology integration and self-efficacy in teaching, such as setting goals and learning experiences. (Ünal et al., 2017).

Under normal conditions of certainty, the success of the implementation and use of an information system can be quantified by indicators such as actual use (how the system and the apps are being used), perceived usefulness, and intent to use (Yoo et al., 2012), or satisfaction. Given the current situation of uncertainty and the mandatory use of e-learning platforms as well as the need to adapt to highly stressful situations, we will consider satisfaction as a key performance indicator of the successful implementation of an information system for the end user. Based on the existing related research we posit that:

- H1: Technology self-efficacy has both direct and mediated effects on the satisfaction with the use of e-learning platforms.
- H2: Technology self-efficacy has both direct and mediated effects on the effective use of e-learning platforms.
- H3: The use of e-learning applications has both a positive direct effect on satisfaction towards the use of e-learning platforms and a mediated effect between technology self-efficacy and satisfaction.

1.3. Technostress

During the pandemic, teachers were exposed to higher levels of stressors due to the need of being always online. Techno-stressors are stimuli, events or demands related to technology, grouped into five categories of ICT stressors in the working environment (Ragu-Nathan et al., 2008), with connected stressors in a non-working environment (Tarafdar et al., 2020). In an academic environment, techno-overload is perceived when teachers are expected to work harder if they use e-learning systems. Techno-invasion appears when academics perceive the borders between private domains and work to be blurred, a situation that was exacerbated in the pandemic when academics taught from their own homes. Techno-complexity occurs when teachers think they are not capable to use tools because they do not have the required/technical skills. This situation seems to appear especially for men, younger people or those with lower computer literacy (Maier et al., 2017).

Techno-insecurity is related to the anxiety of losing one's job due to information systems. It could occur due to over-technologization and socialization of teaching (e.g. the emergence of MOOCs or technology-enhanced learning applications). Techno-uncertainty could be faced when ICT systems are seen as a source of multiple changes in the organization, focusing on the rate at which the software or hardware change, eventually in a digital transformation process, or when teachers feel anxious about the integration of ICT in their teaching activities, thus creating uncertain work expectations or requirements (Li & Wang, 2021).

Technostress consequences at an individual level consist of reduced job satisfaction, productivity, end user satisfaction, and performance (Tarafdar et al., 2020), and increased burnout (Pflügner et al., 2021). Extended working hours are correlated with higher levels of job stress (Jerrim & Sims, 2021; Syvänen et al., 2016) and the need for permanent change within an environment that appears to have grown into a more and more turbulent one, producing progressively stressful working conditions (Fida et al., 2015). Consequently, we posit that:

- H4: Technostress creators have both direct and mediated effects on the satisfaction with the use of e-learning platforms and on their effective use.

As personal factors, organizations could also help employees decrease technostress by implementing support mechanisms like helpdesks or development programs (facilitating conditions) based on employees' perceptions regarding the resource availability that could remove technological barriers (Venkatesh et al., 2012). Tarafdar et al. (2015) found that the presence of inhibiting mechanisms usually improved the use of information systems, resulting in enhanced outcomes: increased productivity or innovation, increased system use-related satisfaction etc.

Technostress inhibitors represent available facilitating resources that could decrease harmful effects caused by technostress creators. Research conceptualized three technostress inhibitors: literacy facilitation, provision of technical support, and facilitation for user involvement in various technology-related decisions. Literacy facilitation (1) is related to plans intending to enhance knowledge and skills (e.g. professional development plans, teamwork, knowledge sharing) (Skaalvik & Skaalvik, 2017). When challenges occur, technical support (2) via various IT support systems could lower the techno-related stress. Involvement facilitation (3) regards the involvement of teachers in the decision-making process related to the implementation of various information systems related to e-learning and research collaboration (Tarafdar et al., 2020). These findings could show that inhibitors, like a high-quality helpdesk, are expected to increase satisfaction with the use of ICT and e-learning systems and to further indirectly affect teachers' inclination to increase the usage of ICT systems at work. We thus expect:

- H5: Technostress inhibitors have both direct and mediated effects on the satisfaction with the use of e-learning platforms and on their effective use.

1.4. Intolerance to uncertainty

Intolerance of uncertainty is the propensity to negatively react in behavioral, emotional, and mental areas when facing uncertain situations (Dugas et al., 2004). Individuals with a higher degree of intolerance to uncertainty would be more relaxed when predictability is possible and, therefore, would intend to decrease or even eliminate uncertainty (Uzun & Karata , 2020).

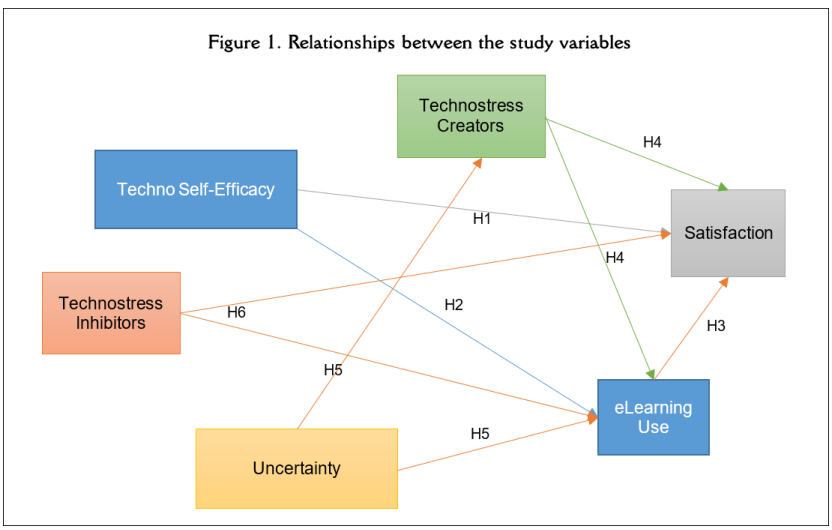
Research on employees' perceived uncertainty at work is scarce. It suggests that perceptions of uncertainty negatively affect job-related satisfaction due to the concern of not being able to foresee what the future will bring, uncertainty being a powerful stressor (Tinaztepe, 2012). It is also possible that individuals with higher intolerance of uncertainty could be more concerned by the uncertainty of further transformations related to the pandemic and its socio-economic consequences (Mertens et al., 2021).

Considering the changing conditions of the current situation, including the suggestions that a state of uncertainty could affect employees' work performance, we posit that:

- H6: The intolerance of uncertainty has indirect effects on the use of e-learning platforms and the satisfaction towards the use of e-learning platforms.

Based on the literature and the above relationships, we assume the existence of the following theoretical model (Figure 1).

Figure 1. Relationships between the study variables



2. Materials and methods

The main objective was to assess the predictive role of technostress creators and inhibitors, techno self-efficacy, and the use of the online platform on teachers' satisfaction in conditions of uncertainty. A quantitative, cross-sectional study was used.

The sample consisted of N=1,517 academics (males 44%) from various Romanian universities from a total of 34,440 individuals, sample error for a confidence level of 95% being +/- 2.5. The sample structure by academic position, fundamental teaching domains, age, and gender are presented in Table 1.

Table 1. Sample characteristics

Academic positions		Fundamental domain		Age (years)	
Assistant	8.2%	Engineering sciences	33.7%	25-35	10.9%
Lecturer	33.4%	Social sciences	26.8%	36-45	34%
Associate Professor	27.7%	Humanities and arts	17.8%	46-55	31.2%
Professor	25.3%	Mathematics and natural sciences	12%	56-60	10.8%
External teacher	5.4%	Biological and biomedical sciences	7.7%	61-65	7.8%
		The science of sports and physical education	2%	Over 65	5.3%

We prepared and sent a set of questionnaires by means of a personalized email: 30,171 email addresses were collected from websites using the open source Abot crawler, then parsing HTML and PDF files for email addresses using regular expressions. The next step consisted of manually cleaning generic email addresses (e.g. contact@, hr@). The results were imported into the survey application. The data was collected in April-May 2021, with a response rate of 5.03%. The participants gave their consent at the beginning of the survey. The study was approved by the council of the Faculty of Psychology and Education Sciences, TRANSYLVANIA University of Brasov, Romania, after checking all the ethical standards for human research studies.

2.1. Instruments

Technostress creators (TC) were measured using the 23-item Technostress scale (Tarafdar et al., 2015). It measures techno-overload (5 items: "I am forced to change my work habits to adapt to new technologies"), Techno-invasion (4 items, "I spend less time with my family due to this technology"), Techno-complexity (5 items, "I often find it too complex to understand and use new technologies"), Techno-insecurity (5 items, "I have to constantly update my skills to avoid being replaced") and techno-uncertainty (4 items, "There are constant changes in the computer software in our organization").

Technostress inhibitors (TI) were measured using the "technostress inhibitors scale"(Ragu-Nathan et al., 2008). The scale conceptualizes three dimensions of technostress: Literacy facilitation (5 items, "Our organization provides end-user training before the introduction of new technology"), Technical support provision (4 items, "The IT department in our organization is well staffed by knowledgeable individuals") and Involvement facilitation (4 items, "We are encouraged to try out new technologies").

The intolerance of uncertainty (IU) was measured using the scale developed by (Carleton et al., 2007). The 12 items are rated on five-point Likert scale measuring responses to uncertainty, ambiguous situations, and the future. They are grouped into two dimensions, Prospective anxiety (7 items; "I can't stand being taken by surprise") and Inhibitory anxiety (5 items; "When it's time to act, uncertainty paralyses me").

Technology self-efficacy (TSE) was measured using the five-item Technology self-efficacy scale (Venkatesh et al., 2012). The 5 items are rated on a five-point Likert scale r and measures the belief in one's ability to successfully perform online tasks in educational settings (e.g., "Whether the use of online technology is difficult or easy"). The use of e-learning platforms in teaching activities (USE) was measured through 8 items created by the authors of this study, measured on a five-point Likert scale. The items were grouped in two dimensions, eee of e-learning platforms in evaluation and monitoring activities (4 items "Monitoring students' progress through evaluation throughout the course") and use of e-learning platforms in teaching activities (4 items: "I use the platform to prepare individualized work tasks").

The satisfaction with the use of e-learning platforms (SAT) was measured through a three-item scale, created by the authors of this study (e.g., "Overall, how satisfied are you with the recent online teaching experience (last month)?"). A ten-point Likert scale was used, ranging from 1 (not at all characteristics of me) to 10 (entirely characteristic of me). Details and results concerning the measurement model and Cronbach's Alpha are provided in the section on Measurement model and online (https://doi.org/10.6084/m9.figshare.19312253).

3. Analysis and findings

The present study employed structural equation modeling, partial least squares (PLS), to analyze data using SmartPLS 3.0 (Ringle et al., 2015). The hypotheses were tested with 5,000 resamples. To analyze and interpret the mediation effects, we used Chin (2010) recommendations. The VIF analysis for evaluating multicollinearity revealed values lower than 3.8, indicating that collinearity is not an issue. All the dimensions were considered reflective while the use of e-learning applications was considered formative.

3.1. Descriptive analysis for the use of online applications and e-learning platforms

The average number of teaching activities in a week for the first semester was 5.26. Most participants (15.9%) stated that they carry out 10 or more activities per week, followed by those who carry out 4 teaching activities per week (15.2%). An in-depth analysis reveals that most professors in teaching positions

(25.6%) spend a long time (between 3-4 hours) preparing an online teaching activity (Table 2). They usually have fewer teaching activities than other teachers, but of a high complexity, which requires careful training. Lecturers are the ones who allocate mostly more than 4 hours a day for the preparation of an activity (24.8%) compared to the associated teachers, who spend less than 1-hour training (11%).

Table 2. Time spent preparing teaching activities for different career levels/teaching positions

		How much time do you spend on average preparing for a 2-hour online teaching activity?					Total
		1-2 hours	2-3 hours	3-4 hours	More than 4 hours	Less than 1 hour	
Assistant	Freq./%	41/32.8%	36/28.8%	17/13.6%	21/16.8%	10/8%	125
Lecturer		142/28%	120/23.7%	99/19.5%	112/22.1%	34/6.7%	507
Associate Professor		118/28.1%	93/22.1%	78/18.6%	104/24.8%	27/6.4%	420
Professor		88/23%	84/21.9%	98/25.6%	87/22.7%	26/6.8%	383
External teacher		27/32.9%	17/20.7%	17/20.7%	12/14.6%	9/11%	82
TOTAL		416/27.4%	350/23.1%	309/20.4%	336/22.1%	106/7%	1,517

An analysis from the perspective of the fundamental teaching domain shows that for exact sciences teachers dedicated more than 4 hours to the preparation of online teaching activities. Between 1 and 2 hours for material, preparation are allocated by teachers in the social sciences fields. Further details regarding preparation time for online teaching activities can be found here: https://doi.org/10.6084/m9.figshare.19312253.

3.2. Measurement model

The model has five constructs with reflective measurements and a construct with formative measurement (the use of e-learning platforms) (Table 3). The composite reliability (CR) (items constantly measure the same construct) for each construct was at least 0.78, showing satisfactory levels of internal consistency.

For assessing convergent validity (how closely a scale is related to other variables or measures of the same construct), the average variance extracted (AVE) for all the constructs were between .550 and .808, satisfying the requirements. The values of indicator loadings for the formative construct were all above .62 indicating adequate convergent validity. Given the good psychometric properties, we kept all the items included in the initial scales.

Table 3. Assessment of the measurement model

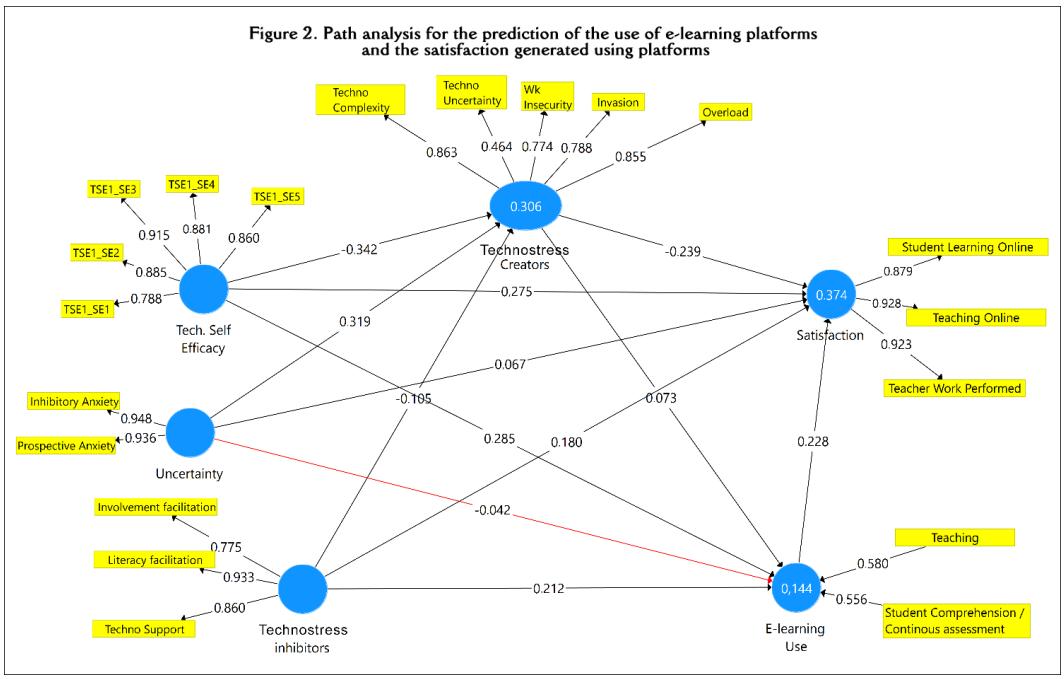
	Cronbach's Alpha	CR	AVE
Intolerance of uncertainty			
Inhibitory anxiety	.848	.892	.623
Prospective anxiety	.795	.859	.552
Technostress inhibitors			
Involvement facilitation	.594	.784	.550
Literacy facilitation	.853	.894	.630
Technical support	.920	.944	.808
Technostress creators			
Complexity	.836	.884	.605
Uncertainty	.774	.868	.687
Insecurity	.791	.856	.546
Invasion	.863	.907	.709
Overload	.859	.898	.642
Technology self-efficacy	.916	.937	.751
Satisfaction towards the use of e-learning	.896	.935	.828

Note. See at https://doi.org/10.6084/m9.figshare.19312253.

3.3. Hypotheses testing

The Pearson correlation (https://doi.org/10.6084/m9.figshare.19312253) showed that TSE was negatively correlated with TC and positively associated with TI, the highest positive association was obtained for SAT, while the associations of TSE with USE was moderate. IU was negatively associated with TI and positively with TC, while the associations with SAT was negative and low. TI correlated positively with the use of e-learning platforms and the satisfaction generated by their use, while the associations of the use dimensions with the TC were negative. To test the hypothesis, we ran several

mediation analyses, using TSE, IU, and TI as antecedents, TC as mediator, and USE and SAT as effects (Figure 2). The total variance explained was 30% for TC, 37% for SAT, and 14% for USE.



The analysis of the direct, indirect, and total effects showed that the relationships between TSE, TI and IU and USE and SAT are mediated by the TC. We found mostly partial mediations, the direct effects between the antecedents and USE and the SAT being significant. We found only one total mediation between the IU and USE through TC. USE does not mediate the association between the IU and SAT, and TC does not mediate the association between TI and the USE (Table 4).

Table 4. Direct, indirect and total effects

Path	Coeff.	t	p	Effect	Hypothesis
TSE→Satisfaction	.275	9.357	<.001	Direct	H1
TSE→TC	-.341	10.914	<.001	Direct	H1
TSE→Use of e-learning	.284	8.970	<.001	Direct	H2
TSE→Use of e-learning→Satisfaction	.064	6.582	<.001	Indirect partial	H3
TC→Use of e-learning	.072	2.279	.022	Direct	H4
Uncertainty→Techno-creators	.318	11.698	<.001	Direct	H6
TI→Satisfaction	.180	7.549	<.001	Direct	H5

Note. See at https://doi.org/10.6084/m9.figshare.19312253.

We have run several multi-group analyses related to gender, academic position and teaching and research domain and none showed significant results.

4. Discussion and conclusions

Most of the study hypothesis were sustained by the data, showing mainly that TC mediates the relationships between TI, TSE and use and satisfaction towards the use of e-Learning platforms. Although the current context is dominated by uncertainty, the hypothesis regarding the direct and indirect effect of uncertainty on the use of the online application in education were partially sustained. In line with previous research, our study showed that TSE is one of the most important antecedents of the use of e-learning platforms and the satisfaction to use them, the direct effects of self-efficacy being positive and significant (Pan, 2020), supporting H1. Other studies showed self-efficacy is a powerful contributing factor related to the ease of use which, combined with behavioral intention, affects the actual use. Our study confirmed previous research, showing that self-efficacy directly impacts the behavioral intention to use technology and actual use (Maican et al., 2019). Our results showed that enhanced self-efficacy is a condition for

technology integration in teaching practices. Regarding the satisfaction with the use of online environments, previous studies confirmed that individuals with high TSE were more satisfied with an online environment, thus self-efficacy could improve satisfaction (Prifti, 2022), supporting H2.

In situations when the mandatory use of a system is studied, the satisfaction of using the system is the most appropriate research performance indicator and not the intention to use it (Chan et al., 2010). Our study showed that the use of e-learning applications has positive direct effects on satisfaction, supporting H3. Furthermore, given the TSE importance, academics' attitudes and involvement are critical for the efficient use of e-learning, leading to increased satisfaction, concluding that usage has a mediator effect between TSE and satisfaction.

It is extensively recognized that self-efficacy is both a supporting factor of work success and a protecting element in stressful working conditions. The negative direct effects of TSE on TC revealed by our results confirm the protective role of self-efficacy against stress (Caprara et al., 2003). High levels of TSE could significantly decrease technostress caused by the complexity of the technology involved in online activities. Academics with higher self-efficacy have higher academic computer-related technology confidence, this belief individuals to prevail over the complexity of the technology, sense of job insecurity, uncertainty, and feelings of techno-invasion in daily academic life (Shu et al., 2011).

Concerning the effects of TC on the use of e-learning, the results showed significant negative effects consistent with a previous study (Upadhyaya & Vrinda, 2021). However, the negative effect of the use of e-learning platforms was smaller than the direct negative effects of TC on user satisfaction. The extensive utilization of ICT in academia could explain the higher levels of technostress which significantly reduce the satisfaction generated by using ICT. Although, previous research focused on general work satisfaction and our results fit into this framework, demonstrating that technostress is responsible for information fatigue, motivation loss, and unhappiness at work (Salah-Eddine & Belaissaoui, 2017). Our results support H4, the partially mediated effect of TSE on USE and SAT, through TC.

TI has a negative direct effect on TC and positive direct effects on USE and SAT. Previous research confirmed the direct effects of TI on job satisfaction and the use of ICT (Upadhyaya & Vrinda, 2021). However, the indirect effects of TI on USE through TC were not significant, albeit marginally, showing that the direct effect of the inhibitors is more important and that this mediation hypothesis is not supported. For SAT, a partial mediation was found, given both the significant direct and indirect effects (through technocreators) of TI. The direct and serial indirect effects of TI on satisfaction (through TC and USE) were also significant. We conclude that TI reduces the effects of technostress (Ragu-Nathan et al., 2008), sustaining organizational mechanisms and adjustment to reduce the negative outcomes of increasing ICT use and thus, explaining a higher satisfaction of using ICT in education (Jena, 2015) (H5 partially supported).

The pandemic has affected billions of people, and the uncertainty of tomorrow has important consequences on individuals' behavior. Uncertainty involves increased anticipation of a negative situation making people less capable of coping with negative events, which can explain higher levels of perceived stress and maladaptive behaviors given the increased anticipation of a negative situation (Grupe & Nitschke, 2013). Stress and uncertainty have left their mark on the work of teachers and students alike. The need to move all activities online created a new context, emergency remote teaching (Hodges et al., 2020). Students and staff were forced to manage technical concerns, spatial arrangements, family/conflicting responsibilities, and physical and mental health issues. Our results confirmed the direct positive effects of uncertainty on technostress creators, showing that uncertainty could explain the increase of technostress.

However, TC had a weak direct positive association with USE, while the direct effect of uncertainty on USE was not significant. The effects of uncertainty on satisfaction towards the use of e-learning were negative, both direct and indirect, showing that satisfaction is connected with uncertainty, even when technostress is involved. Although we expected direct effects between uncertainty and USE, these effects did not occur, perhaps because the uncertainty situation can be seen as a global one, with medium and long-term effects. As employees, teachers did not feel the short-term effects, since they know what skills and tools to use for carrying out their activities. Given that online/remote teaching was the only available option, the relationship between technostress, platform use and uncertainty cannot be very clear, concluding that

H6 is partially supported. While self-efficacy is a protective factor against stress, uncertainty is an aversive state, explaining low levels of satisfaction with the use of technology. Although research focused on the negative associations between uncertainty and job satisfaction (Bordia et al., 2004), we believe our results could also provide empirical evidence for the harmful effects of uncertainty.

The multi-group analysis did not show differences between the analyzed groups (teaching positions, gender, teaching/research field). Previous research found differences in technostress concerning gender and work experience (Marchiori et al., 2019; Ragu-Nathan et al., 2008). Our study did not reveal significant group differences, the results being convergent with recent studies (Li & Wang, 2021). Although we expected differences, the pandemic context probably canceled them out.

Although the pandemic context is fraught with uncertainty in explaining the high level of stress, its negative weight can be counteracted by a high level of TSE. Equally, the support provided by organizations through techno-inhibitors can be a major factor, sustaining the use of online technologies, even under normal conditions. Even if the universities are promoters of advanced research, academics have lifelong learning expectations, even in the case of "basic" tools such as an e-learning platform, especially related to the content which requires a significant effort to transpose for online teaching (i.e., from the platform we want much more than simply sharing presentations). Also, finding solutions for technical disciplines involving special equipment (e.g., remote-virtual-laboratories) also requires training.

This study may contribute to current research on self-efficacy, e-learning use and technostress in uncertain situations. First, TSE is under researched in connection with e-learning use satisfaction, even if self-efficacy is an important factor of task performance. Second, research regarding intolerance to uncertainty among teachers in the context of e-learning is extremely few. Although uncertainty is characteristic of the pandemic, as a factor resulting from the person-situation interaction, uncertainty is not manifested directly over the satisfaction of using e-learning, but only through factors related to the use of technology. Therefore, uncertainty is an indirect "influencer" in certain situations where the state of uncertainty caused by various macrosocial events remains high (e.g., conflicts, environmental issues, various etc.). Third, technostress creators and technostress inhibitors play an important role in satisfaction, confirming the correctness of our choice. A possible implication could result in providing not only technical support but also training programs for teachers to develop their self-regulation and coping strategies in uncertain situations. Forth, we developed a tool to identify teachers' use of e-learning platforms together with a tool for identifying teaching satisfaction. Such tools can be used in the future to evaluate the efficiency of teaching-learning from the students' perspective. These could also be used as self-assessment tools for university teachers to improve their online teaching practices.

The pandemic has marked an essential point in using IT technologies in learning, many once-regular-teachers will use them more intensely in the future, being as an impetus towards digital transformation into HEIs. A possible limitation could refer to the challenging moment when the study was run, with universities struggling to find quick solutions to continue their work. As expected, not all the chosen solutions were the best, since each Romanian university (or even teacher) decided which online platform to use for mandatory online learning, implying differences in system complexity, ease of use, perceived usefulness etc. Another limitation is that some universities have already implemented functional e-learning systems, but which were designed and used on a much smaller scale. Thus, some teachers were already familiar with the e-learning systems, at least partially, and so the pandemic was not as challenging as for those who had not used these systems in the past and had to adapt on the fly.

The lack of data on the actual use of the platforms could be another limitation, as only statistics related to the download of certain applications are public but not information on their use. This limitation could be solved in a future study applied at the level of a single HEI that would allow access to the usage logs of the applications from which the current usage would be extracted. The lack of actual user data and the use of only self-reported values could also cause biased data analysis. In line with this, a possible self-selection bias could be added, given that the participants volunteered to participate in this study, and their participation could have been influenced by their great interest in technology and its use. In addition, the data was collected during the pandemic, when the topic of e-learning was extensively discussed. Another limitation of this study was the cross-sectional nature of the data used, making it impossible to draw causal

inferences. A future direction of study is to identify the degree of use of the most popular platforms as well as the comparative study of the use of e-learning technology at the teacher level after the COVID pandemic.

Authors' Contribution

Idea, A.M.C, C.I.M; Literature review (state of the art), C.I.M; Methodology, A.M.C; Data analysis, A.M.C, C.I.M; Results, A.M.C, C.I.M; Discussion and conclusions, A.M.C, C.I.M; Writing (original draft), C.I.M; Final revisions, A.M.C, C.I.M; Project design and funding agency, A.M.C.

Funding Agency

This study was supported by the TRANSYLVANIA University of Brasov, Romania.

References

Al-Samarraie, H., Teng, B.K., Alzahrani, A.I., & Alalwan, N. (2018). E-learning continuance satisfaction in higher education: A unified perspective from instructors and students. *Studies in Higher Education*, *43*, 2003-2019. https://doi.org/10.1080/03075079.2017.1298088

Almuwais, A., Alqabbani, S., Benajiba, N., & Almoayad, F. (2021). An Emergency Shift to e-Learning in Health Professions Education: A Comparative Study of Perspectives between Students and Instructors. *International Journal of Learning, Teaching and Educational Research*, *20*(6). https://doi.org/10.26803/ijlter.20.6.2

Bakar, N.S.A., Maat, S.M., & Rosli, R. (2018). A systematic review of teacher's self-efficacy and technology integration. *International Journal of Academic Research in Business and Social Sciences*, *8*(8). https://doi.org/10.6007/IJARBSS/V8-I8/4611

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W H Freeman/Times Books/ Henry Holt & Co. https://bit.ly/2FVgMwH

Bordia, P., Hobman, E., Jones, E., Gallois, C., & Callan, V.J. (2004). Uncertainty during organizational change: Types, consequences, and management strategies. *Journal of Business and Psychology*, *18*(4), 507-532. https://doi.org/10.1023/B:JOBU.0000028449.99127.f7

Caprara, G.V., Barbaranelli, C., Borgogni, L., & Steca, P. (2003). Efficacy beliefs as determinants of teachers' job satisfaction. *Journal of Educational Psychology*, *95*(4), 821-832. https://doi.org/10.1037/0022-0663.95.4.821

Carleton, R.N., Norton, M.A.P.J., & Asmundson, G.J.G. (2007). Fearing the unknown: A short version of the Intolerance of Uncertainty Scale. *Journal of Anxiety Disorders*, *21*(1), 105-117. https://doi.org/10.1016/j.janxdis.2006.03.014

Casero-Béjar, M.O., & Sánchez-Vera, M.M. (2022). Cambio de modalidad presencial a virtual durante el confinamiento por COVID-19: Percepciones del alumnado universitario. *RIED*, *25*, 243-260. https://doi.org/10.5944/ried.25.1.30623

Chan, F.K.Y., Thong, J.Y.L., Venkatesh, V., Brown, S.A., Hu, P.J.H., & Tam, K.Y. (2010). Modeling citizen satisfaction with mandatory adoption of an E-Government technology. *Journal of the Association for Information Systems*, *11*(10), 519-549. https://doi.org/10.17705/1jais.00239

Chin, W.W. (2010). How to write up and report PLS analyses. In V. Esposito-Vinzi., W. Chin, J. Henseler, & H. Wang. (Eds.), *Handbook of Partial Least Squares* (pp. 655-690). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-32827-8_29

Dugas, M.J., Buhr, K., & Ladouceur, R. (2004). The role of intolerance of uncertainty in etiology and maintenance. In R. G. Heimberg, C. L. Turk, & D. S. Mennin (Eds.), *Generalized anxiety disorder: Advances in research and practice* (pp. 143-163). The Guilford Press. https://bit.ly/3cXDWFt

Fida, R., Paciello, M., Tramontano, C., Barbaranelli, C., & Farnese, M.L. (2015). Yes, I can": The protective role of personal self-efficacy in hindering counterproductive work behavior under stressful conditions. *Anxiety, Stress, & Coping*, *28*(5), 479-499. https://doi.org/10.1080/10615806.2014.969718

Grupe, D.W., & Nitschke, J.B. (2013). Uncertainty and anticipation in anxiety: An integrated neurobiological and psychological perspective. *Nature Reviews Neuroscience*, *14*(7), 488-501. https://doi.org/10.1038/nrn3524

Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020). *The difference between emergency remote teaching and online learning*. EDUCAUSE Review. https://bit.ly/3bk6MPI

Jena, R.K. (2015). Impact of technostress on job satisfaction: An empirical study among Indian academician. *The International Technology Management Review*, *5*(3), 117-124. https://doi.org/10.2991/itmr.2015.5.3.1

Jerrim, J., & Sims, S. (2021). When is high workload bad for teacher wellbeing? Accounting for the non-linear contribution of specific teaching tasks. *Teaching and Teacher Education*, *105*, 103395-103395. https://doi.org/10.1016/j.tate.2021.103395

Li, L., & Wang, X. (2021). Technostress inhibitors and creators and their impacts on university teachers. *Cognition, Technology and Work*, *23*, 315-330. https://doi.org/10.1007/s10111-020-00625-0

Maican, C.I., Cazan, A.M., Lixandroiu, R.C., & Dovleac, L. (2019). A study on academic staff personality and technology acceptance: The case of communication and collaboration applications. *Computers and Education*, *128*, 113-131. https://doi.org/10.1016/j.compedu.2018.09.010

Maier, C., Laumer, S., Eckhardt, A., & Weitzel, T. (2017). Giving too much social support: Social overload on social networking sites. *European Journal of Information Systems*, *24*(5), 447-464. https://doi.org/10.1057/ejis.2014.3

Marasi, S., Jones, B., & Parker, J.M. (2022). Faculty satisfaction with online teaching: A comprehensive study with American faculty. *Studies in Higher Education*, *47*(3), 513-525. https://doi.org/10.1080/03075079.2020.1767050

Marchiori, D.M., Mainardes, E.W., & Rodrigues, R.G. (2019). Do Individual characteristics influence the types of technostress reported by workers? *International Journal of Human-Computer Interaction*, *35*(3), 218-230. https://doi.org/10.1080/10447318.2018.1449713

Mertens, G., Duijndam, S., Smeets, T., & Lodder, P. (2021). The latent and item structure of COVID-19 fear: A comparison of four COVID-19 fear questionnaires using SEM and network analyses. *Journal of Anxiety Disorders*, *81*. https://doi.org/10.1016/j.janxdis.2021.102415

Mouakket, S., & Bettayeb, A.M. (2015). Investigating the factors influencing continuance usage intention of Learning management systems by university instructors: The Blackboard system case. *International Journal of Web Information Systems*, *11*(4), 491-509. https://doi.org/10.1108/IJWIS-03-2015-0008

Pan, X. (2020). Technology acceptance, technological self-efficacy, and attitude toward technology-based self-directed learning: Learning motivation as a mediator. *Frontiers in Psychology*, *11*, 2791-2791. https://doi.org/10.3389/fpsyg.2020.564294

Pflügner, K., Maier, C., Mattke, J., & Weitzel, T. (2021). Personality Profiles that put users at risk of perceiving technostress: A qualitative comparative analysis with the big five personality traits. *Business and Information Systems Engineering*, *63*(4), 389-402. https://doi.org/10.1007/s12599-020-00668-7

Polly, D., Martin, F., & Guilbaud, T.C. (2021). Examining barriers and desired supports to increase faculty members' use of digital technologies: Perspectives of faculty, staff and administrators. *Journal of Computing in Higher Education*, *33*(1), 135-156. https://doi.org/10.1007/s12528-020-09259-7

Pressley, T., & Ha, C. (2021). Teaching during a pandemic: United States teachers' self-efficacy during COVID-19. *Teaching and Teacher Education*, *106*, 103465-103465. https://doi.org/10.1016/j.tate.2021.103465

Prifti, R. (2022). Self-efficacy and student satisfaction in the context of blended learning courses. *Open Learning: The Journal of Open, Distance and e-Learning*, *37*, 111-125. https://doi.org/10.1080/02680513.2020.1755642

Ragu-Nathan, T.S., Tarafdar, M., Ragu-Nathan, B.S., & Tu, Q. (2008). The consequences of technostress for end users in organizations: Conceptual development and validation. *Information Systems Research*, *19*(4), 417-433. https://doi.org/10.1287/isre.1070.0165

Ringle, C., Wende, S., & Becker, J. (2015). *SmartPLS 3*. SmartPLS GmbH. https://bit.ly/3vxn0w2

Rokhimah, S., & Sirait, A. (2021). Educator satisfaction using LMS with ICT infrastructure as a mediation variable. *RSF Conference Series: Business, Management and Social Sciences*, *1*(3), 81-87. https://doi.org/10.31098/bmss.v1i3.291

Salah-Eddine, M., & Belaissaoui, M. (2017). Technostress, coping and job satisfaction model of information systems. In *Proceedings - 2016 International Conference on Computational Science and Computational Intelligence, CSCI 2016* (pp. 139-142). IEEE. https://doi.org/10.1109/CSCI.2016.0033

Shu, Q., Tu, Q., & Wang, K. (2011). The impact of computer self-efficacy and technology dependence on computer-related technostress: A social cognitive theory perspective. *International Journal of Human-Computer Interaction*, *27*(10), 923-939. https://doi.org/10.1080/10447318.2011.555313

Skaalvik, E.M., & Skaalvik, S. (2017). Still motivated to teach? A study of school context variables, stress and job satisfaction among teachers in senior high school. *Social Psychology of Education*, *20*(1), 15-37. https://doi.org/10.1007/s11218-016-9363-9

Stickney, L.T., Bento, R.F., Aggarwal, A., & Adlakha, V. (2019). Online Higher Education: Faculty Satisfaction and Its Antecedents. *Journal of Management Education*, *43*(5), 509-542. https://doi.org/10.1177/1052562919845022

Syvänen, A., Mäkiniemi, J.P., Syrjä, S., Heikkilä-Tammi, K., & Viteli, J. (2016). When does the educational use of ICT become a source of technostress for Finnish teachers? *Seminar.Net*, *12*(2), 95-109. https://doi.org/10.7577/seminar.2281

Tarafdar, M., Maier, C., Laumer, S., & Weitzel, T. (2020). Explaining the link between technostress and technology addiction for social networking sites: A study of distraction as a coping behavior. *Information Systems Journal*, *30*(1), 96-124. https://doi.org/10.1111/isj.12253

Tarafdar, M., Pullins, E., Bolman, & Ragu-Nathan, T.S. (2015). Technostress: Negative effect on performance and possible mitigations. *Information Systems Journal*, *25*(2), 103-132. https://doi.org/10.1111/isj.12042

Tinaztepe, C. (2012). The effect of desire for change on the relationship between perceived uncertainty and job related affective well being. *International Journal of Social Sciences and Humanity Studies*, *4*(2), 127-136. https://bit.ly/3OPV1yv

Ünal, E., Yamaç, A., & Uzun, A.M. (2017). The Effect of the teaching practice course on pre-service elementary teachers' technology integration self-efficacy. *Malaysian Online Journal of Educational Technology*, *5*(3), 39-53. https://bit.ly/3zWwcg9

UNESCO (Ed.) (2020). *COVID-19 response–remote learning strategy. Remote learning strategy as a key element in ensuring continued learning*. UNESCO. https://bit.ly/3JmFlfn

Upadhyaya, P., & Vrinda (2021). Impact of technostress on academic productivity of university students. *Education and Information Technologies*, *26*, 1647-1664. https://doi.org/10.1007/s10639-020-10319-9

Uzun, K., & Karata , Z. (2020). Predictors of academic self efficacy: Intolerance of uncertainty, positive beliefs about worry and academic locus of control. *International Education Studies*, *13*(6), 104-116. https://doi.org/10.5539/ies.v13n6p104

Venkatesh, V., Thong, J.Y.L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use technology. *MIS Quarterly*, *36*(1), 157-178. https://doi.org/10.2307/41410412

Yoo, S.J., Huang, W.H., & Lee, D.Y. (2012). The impact of employee's perception of organizational climate on their technology acceptance toward e-learning in South Korea. *Knowledge Management & E-Learning: An International Journal*, *4*(3), 359-378. https://doi.org/10.34105/j.kmel.2012.04.028

Comunicar, n. 74, v. XXXI, 2023 | Media Education Research Journal | ISSN: 1134-3478; e-ISSN: 1988-3478
www.comunicarjournal.com

Political empowerment among young voters: Social media, partisanship and the moderating role of political interest

Empoderamiento político entre jóvenes votantes: Redes sociales, partidismo y papel moderador del interés político

Rehan Tariq. PhD Student, School of Communication, Universiti Sains Malysia, Pulau Pinang (Malaysia) (rehantariq@gcuf.edu.pk) (https://orcid.org/0000-0001-7028-4430)
Dr. Fakhta Zeib. Lecturer, Department of Mass Communication, Government College University, Faisalabad (Pakistan) (zeib@staff.uni-marburg.de) (https://orcid.org/0000-0002-6889-2995)

ABSTRACT

Despite recent efforts to examine the political outcomes of social media use, little is known about the reinforcement of political empowerment and moderating effect of political interest. It is vital to understand the elements that influence the level of political empowerment. Therefore, this paper used a sample of Pakistani voters (n=410) aged 18-29 to offer insight into how social media use alongside other political behaviors, such as partisanship, political expression, and political interest, affect political empowerment. Social Identity Theory (SIT) and rational choice theory provided the theoretical underpinning for the variables of this study. The statistical analyses were performed using Partial Least Squares (PLS) to assess the effects of four variables i.e., social media use, political partisanship, political expression, and political interest, on political empowerment. This study made a significant contribution to the research literature by combining SIT and rational choice theory in one framework. This study also expands the literature on political interest by introducing political interest as a moderator between the inconsistent relationship of political expression and political empowerment. Our results demonstrated that partisanship and social media use positively influenced political expression among young voters. Moreover, political interest positively moderated the relationship between political expression and political empowerment.

RESUMEN

A pesar de los recientes esfuerzos para examinar los resultados políticos del uso de las redes sociales, se sabe poco sobre el refuerzo del empoderamiento político. Es vital comprender los elementos que influyen en el nivel de empoderamiento político. Por lo tanto, este documento utilizó una muestra de votantes paquistaníes (n=410) del grupo de edad de 18 a 29 años para ofrecer información sobre cómo el uso de las redes sociales junto con otros comportamientos políticos, como el partidismo, la expresión política y el interés político, afectan al empoderamiento político. La Teoría de la Identidad Social (TIS) y la Teoría de la Elección Racional proporcionaron la base teórica para las variables de este estudio. Los análisis estadísticos se realizaron utilizando mínimos cuadrados parciales (PLS) para evaluar los efectos de cuatro variables, es decir, el uso de las redes sociales, el partidismo político, la expresión política y el interés político, en el empoderamiento político. Este estudio hizo una contribución significativa a la literatura de investigación al combinar ambas teorías en un solo marco. El trabajo también amplía la literatura sobre el interés político al introducirlo como moderador entre la relación inconsistente de la expresión política y el empoderamiento político. Los resultados demostraron que el partidismo y el uso de las redes sociales influyeron positivamente en la expresión política entre los votantes jóvenes. Además, el interés político moderó positivamente la relación entre la expresión política y el empoderamiento político.

KEYWORDS | PALABRAS CLAVE

Young voters, political empowerment, social media, partisanship, political interest, political expression.
Votantes jóvenes, empoderamiento político, redes sociales, partidismo, interés político, expresión política.

Received: 2022-03-10 | Reviewed: 2022-04-26 | Accepted: 2022-07-15 | OnlineFirst: 2022-10-30 | Published: 2023-01-01
DOI https://doi.org/10.3916/C74-2023-08 | Pages: 95-106

1. Introduction

The feasibility of social media's political discourse in participatory democracy, which leads to democratic political actions, has been under investigation for many years. Apart from theoretical work, many scholars have also attempted to lay out an empirical foundation of the practicability of social media as a facilitator of public democratic discourse (Boulianne, 2015). However, the position of social media political discourse is stronger in comparison to traditional media political discourse i.e., television, and its role in sustaining democratic political values in the social and political context of Pakistan. For instance, many social and political influences and pressures put barriers to the coverage of news on traditional media for fulfilling their sensitive role in sustaining democracy in Pakistan. Television channels in Pakistan are run in two different setups. Pakistan's national television is run by the state and private television channels are run under private media ownership. Being the mouthpiece of the government, Pakistan's national media, particularly, broadcast limited independent rational debate, whereas, in the case of private media, the situation is a bit different. Private media is controlled mainly by advertising, government, and cross-media ownership (Riaz, 2007).

Such mediated discourse does not serve democratic political values and does not encourage the metaphor of transparency. Moreover, several years of multiple dictatorship regimes and the limited scope of freedom of expression have pushed the country into political and economic turmoil. For the rest of the period, Pakistan faced barriers in the execution of the political process and the smooth transfer of democratic regimes from one political party to another. However, during and after the 2013 election, the transformation in the political structure of Pakistan provided ample space for all existing and emerging political parties. Pakistan's political structure was based on two-party politics (Pakistan Muslim League and Pakistan People's Party) for decades. However, the emergence of a third political party, Pakistan Tehreek-e-Insaf, transformed the political structure in Pakistan. This transformation resulted in the expansion of not only the number of political parties, but also in the expansion within the political parties has also been witnessed. All political parties' social media cells were functional before the 2013 election, which reduced their dependency on traditional media for the coverage of their political campaigns. Political parties, their supporters, and voters are extensively using social media platforms for their political expression (Zeib, 2022). Nevertheless, the emergence of new information and communication technologies calls for a re-examination of traditional participatory culture and new spaces for public discussion to encourage democratic politics.

1.1. Antecedents of political empowerment

The study is based on two arguments about the relationship of political expression and political empowerment. First, the study measures the effects of social media use and political partisanship on political expression, and second, that political expression enhances political empowerment among youth under the condition of political interest as a moderating factor. Indeed, we found strong evidence that social media political expression is a very substantial predictor of the political empowerment of the users (Loader et al., 2014). However, the question about the strength of this relationship and under what conditions it exists has great importance in a democratic structure of any country. The political system of Pakistan is potentially influenced by voters' strong political affiliations. The general voters of Pakistan continue to show their affiliations with political parties for democratic political participation, and the other side of the coin is that political parties also strengthen their party base with their party supporters and voters (Tariq et al., 2022).

There is an increasing academic interest in exploring how dynamics of political communication change when communicated by a politically interested self. Those with a higher level of political interest are likely to produce more political empowerment than those with a lower level of political interest or having no political interest at all (Maurissen, 2020). Given the above-mentioned proposition, the scholarship discussed in the study is based on an inconsistent relationship between political expression and the political empowerment of the users. Nevertheless, the main purpose of this study is to investigate the inconsistent relationship between political expression and political empowerment, which compelled the researchers to explore a political interest, possibly enhancing the efficacy of political expression for political empowerment.

Nevertheless, this ongoing study is designed to know whether social media political expression and partisan-based political expression enhance youngsters' political empowerment and further, whether increased political empowerment is subject to the moderating effects of political interest. However, we propose the following hypotheses.

- H1: Social media use has a direct positive link with the political expression of youngsters.
- H2: Partisanship has a direct positive link with the political expression of youngsters.
- H3: Political expression has a direct positive link with political empowerment of youngsters.
- H4: The relationship between political expression and political empowerment is positively moderated by political interest.

2. Theoretical underpinning

The rational choice theory posits that individual's economic and social behaviors are based on their self-interests. People select from many options and make a deliberate choice by making cost and benefit comparisons, rather than what psychologists state the unconscious or semi-rational decisions. Many sociologists adapted the rational choice theory to explain social exchanges i.e., the calculation of costs and rewards in social relationships drive social behaviors (Blau, 1964; Homans, 1958). Nevertheless, Harrop and Miller (1987) explain the phenomenon of rational choice in partisanship and argue that partisanship influences the voting behavior of an individual when party affiliation, which is already based on an individual's self-interest, is in accordance with the policy proposals of a political party. They put the bases of political partisanship on the rational choices of voters. They further explain that people make deliberate decisions based on rational choices, having a justification behind every decision. In this framework, we believe that in a democratic structure, political values such as liberal and independent political thoughts and independence in making choices in the decision-making (rational choices) process should also prevail in the political system of the country, which is basically argued by rational choice theorists.

Citizens' capability to evaluate the democratic quality of the political process is a determinant of political interest. On the one hand, if deliberation enhances the political empowerment of citizens, it also requires political interest in the deliberation process (Morrell, 2005). Therefore, we expect that if citizens have more political interest, they will exhibit more political empowerment by making rational choices.

The theoretical foundation of the phenomenon of partisanship is established on Henry Tajfil's Social Identity Theory (SIT) developed in 1974. The theory explains the behavioral aspects of individuals' group belonging (Tajfel, 1974). SIT explains the cognitive process of the inner self of an individual and the resulting related behavioral motivation in a group. In the democratic political structure of any country, people want to be identified with a certain political party affiliation. The desire to be attached to different social and political groups is a human instinct. Partisanship possesses a very significant place in a person's political life that continues to evolve stronger in a youngster's life (Shively, 1979). Because of party loyalties, the impact of political partisanship is very deep as it transmits from one generation to the next.

Taking up the position from the processes identified by SIT, political partisanship has the enforcement value that motivates the partisans to behave in a particular way associated with political groups. Although the members of political groups are mostly bounded by ideological associations rather than proximal attachment i.e., they share common views and interests in their group circle. On the other hand, the members, who have some ideological affiliation, try to seek people with the same traits, behaviors, and viewpoints that are found in in-group partisans and encourage the other members to update their political knowledge (Großer & Schram, 2006).

3. Literature review

The relationship of social media political expression with political efficacy attained considerable research attention from communication scholars throughout the world. More recent scholarship is focusing on the political discourse of social media for youngsters' political awareness and empowerment, news seeking, political campaigning, and online and offline political activities (Ahmad et al., 2020), some of them using cross-sectional and some using panel data.

However, during the process of a systematic review of relevant literature, we pondered on the previous and recent research trends in media, communication, and political science studies. The scholarship on this topic is supported by a plethora of studies based on three theoretical assumptions. The first assumption deals with the relationship of social media and partisan-based political expression with political empowerment. The second is about the causal but inconsistent relationship between political expression and political empowerment, and the third, is about the moderating variables in the relationship of political expression and political empowerment.

3.1. Social media use encourages political expression

Social media use refers to the purposive and frequent social media use to get political information (Zolkepli et al., 2018). We connected this with political information to develop an understanding of political issues. Popular social media websites have started a new arena of discussions, talks, and other forms of expression. The political expression refers to expressing personal emotions and feelings about politics through a variety of social media tools (Chen & Chan, 2017). The ability of social media to enhance the predictable outcomes of individual expression by communicating it to an unlimited network of members immediately has established its unique place in communication research (Gil-de-Zúñiga et al., 2014). In addition, the younger generation tend to use social media more consistently for active participation and social integration (Colás-Bravo et al., 2013). The fast-growing technological changes in digital media have opened opportunities of political expression for youngsters in Pakistan as well. Almost all political parties and political candidates have social media accounts, sharing several posts and tweets in a day. Moreover, young voters are making substantial use of social media in seeking political information and discussing and sharing political views (Ali & Fatima, 2016). Skoric et al., (2016) initiated an empirical investigation on the role of social media use in encouraging the political expression of users in Confucian Asia. They found that social media use encourages users to share their political selves and to express their views on social media.

3.2. Political partisanship encourages political expression

Research depicts that those having strong partisanship or politically extreme ideological views tend to express themselves politically more via various online and offline platforms than those with a low level of partisanship or with non-partisanship (Moffett & Rice, 2018). Partisanship refers to the sense of closeness and attachment of an individual towards a particular political party (Huddy et al., 2010). To be identified with a political party shows a strong sense of psychological attachment to that political party. Furthermore, such expressions associated with party identifications are positively related to an increase in individuals' political empowerment and motivation to participate in political activities (Dancey & Goren, 2010). However, weak or moderate partisans were less likely to express their minority views on Facebook because of the difference of opinion with those who are in the majority (Kim, 2018). Given the moderation effects of higher agreement among cohorts and higher reliance on political parties, strong political partisanship also enhances cognitive and behavioral engagement in political discussions and various forms of online expression (Chan, 2018).

3.3. Inconsistent relationship between political expression and political empowerment

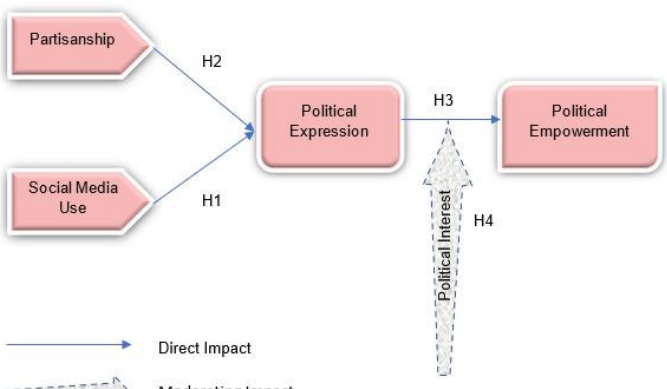
So far, the academic literature is fluctuating between an overly optimized view of social media's political effects and a critical evaluation of the use of these platforms. There is a sizeable body of research that provides evidence that the effects of social media use on political empowerment are uneven and hardly symmetrical, and that the effects vary depending on the context, intervening factors, demographics, etc. Political empowerment is a psychological phenomenon (Spreitzer, 1995) that refers to citizens' ability to analyze the political issues, make political decisions and exercise their capacity to organize and mobilize the community freely and independently (van-Dop et al., 2016). According to Boulianne (2015), the effects of social media political expression vary among youngsters and the general population, different types of users, active and passive social media use, interest-based uses, and cross-sectional and panel data. However, a comprehensive examination of the aforementioned research divulges that substantial gaps exist within the literature, specifically regarding the use of intervening variables, which may help in

enhancing the effect of the explanatory variable. So far, we have concluded that the scientific community has a shared consensus on the relationship of political expression and its effects on political empowerment of users. However, confirmation about the strength of this relationship i.e., strong or weak, direction of relationship i.e., positive or negative, and whether it is a direct or indirect relationship is not yet clear (Boulianne, 2015). Many studies, despite having a positive relationship between political expression and political empowerment, could not explain the high variance (Gil-de-Zúñiga et al., 2014). The central question here is whether the effects are primarily related to digital media use or associated with the people who are more politically interested. This argument firstly shows an inconsistent or triggering relationship of political expressions with political empowerment on which many researchers agree and secondly, it explains the importance of the moderation of political interest in this inconsistent relationship.

Political interest refers to personal attention towards political news and current affairs with keen observation of political activities and seeking political information from everyday communication (Himelboim et al., 2012). Political interest is contested by many contemporary communication scholars as a key for sustaining democracy and it has been found that political interest is a decisive precursor for political efficacy i.e., self-understanding of political abilities and that one's political actions affect the political process (Prior, 2019). Kalogeropoulos et al. (2017) found a positive spiral of reinforcement and suggest that highly motivated or politically active citizens are more likely to get into lengthy discussions and comment on others' posts, which primarily requires political interest to process the information. Our scholarship implies that if political interest can prove to be an effective explanatory variable, it also qualifies as an effective moderating variable.

However, this article builds on the question: "Can the individual's political interest enhance the capacity of political expression for political empowerment of youngsters?", by specifying that effective political expression is expected to have a strong relationship with political empowerment, which is moderated by political interest to achieve the desired goals. Nevertheless, Bimber et al. (2015) revealed contradictory arguments and found a positive and consistent relationship of digital media use with political talks and voting, for having lower political interest. However, we believe a negative relationship exhibits an equally effective implication of political interest as an important moderating variable.

Figure 1. Conceptual framework



Hence, on par with the previously mentioned arguments and in the light of previous literature, the viability of social media or partisan-based political expression for the models of political empowerment weakens, declining the direct paradigm of political expression and political empowerment and engagement. Consequently, we found a strong agreement on a substantial moderating variable of political interest as an important legitimizing mechanism. Based on the previous literature and hypotheses, we design the following conceptual framework as shown in Figure 1.

4. Methods and measures

4.1. Study setting and design

Because of the increasing importance of youth in Pakistan, we targeted young voters as the unit of analysis in this study. According to the ministry of youth affairs in Pakistan, people aged 15 to 29 are considered youth (Ashraf et al., 2013). However, the voting age in Pakistan starts from 18 years old, therefore, individuals aged 18 to 29 years were considered as the study population. The most recent population census reported that the total population of Pakistan is 207,774,520. The young voters comprise 29% of the total population which is approximately 60,254,611 (United Nations Development Pakistan, 2018).

To determine the sample size and demonstrate confident findings, we combined G-power analysis with Krejcie and Morgan (1970) formula. G-power 3.1.9.7 determined a sample size of 85 for four predictors. A power of 0.80 is calculated for this sample size. As the study population was heterogenous, to yield maximum power it is recommended to increase the sample size (Hair et al., 2019). Here, we applied the Krejcie and Morgan (1970) formula and increased the sample size to 385. We received 410 usable responses that yield a power of 0.9, which is sufficient to claim confident findings.

Due to the COVID-19 spread, we collected the data through an online survey using snowball sampling that is appropriate in the absence of a sample frame (Bobbie, 2013). The survey took four months to complete; from July to October 2020. Sarstedt et al. (2019) reported that snowball sampling produces sample bias and generalizability problems.

To address these issues, increasing the sample size and comparing the gender ratio in sample demographics with the ratio of national population statistics are recommended. G-power analysis, the male and female ratio in demographic data, demonstrates the true representation of the population of Pakistan (Table 1). Hence, the issues of true representation and sample bias were addressed.

4.2. Measurement

The responses of participants were recorded on a 5-point Likert scale. For social media use, political interest and political empowerment scale were used ranging from "strongly disagree" to "strongly agree". However, for the partisanship scale, it ranged from "weak" to "extremely strong", and for political expression from "never" to "all the time". The social media use items were adapted from the research of Zolkepli et al. (2018), political interest items from Bimber et al. (2015), Becker and Copeland (2016) and Himelboim et al. (2012), political empowerment items from van-Dop et al. (2016) and Spreitzer (1995), partisanship items from Huddy et al. (2010) and Chan (2018), and political expression items from Gil-de-Zúñiga et al. (2014) and Chen and Chan (2017).

Though all the items used to measure variables were adapted from past research, still we went through a rigorous validation process. An expert panel comprising three academics was consulted. Their feedback was incorporated to modify the items. After this, we conducted a pilot study among 50 respondents. The results of the pilot study assisted us in refining the measurement scale for this study.

5. Results

5.1. Demographics

The demographics show that most of the participants were male (52.20%), followed by female (47.80%). The statistics on the world's social media users back up these results (Barnhart, 2021).

The largest group (43.20%) of respondents were aged between 22 and 25, which coincides with the previous study, showing that the highest percentage of young social media users were between the age of 21 to 25 years in Pakistan (Zulqarnain & Taimur-ul-Hassan, 2017). Table 1 presents demographic details as well as interesting characteristics of the Pakistani population published by the government of Pakistan referring to our findings.

Table1. Demographics (n=410)			
	Description	**Percent (%)**	**National Population (%) (Pakistan Bureau of Statistics, 2021)**
Gender	Male	52.20	51.02
	Female	47.80	48.98
Age	18-21	21.70	18 to 29 Years 29
	22-25	43.20	
	26-29	35.10	
Education Level	Less than high school	1.20	Literacy Rate 60
	High School	8.30	
	Madrassa	1.50	
	Bachelor's Degree	47.10	
	Master's Degree	39.30	
	PhD Degree	1.00	
	Prefer not to answer	1.70	
Profession	Full Time	32.00	Labor Force Unemployment 6.30
	Part Time	6.10	
	Self-Employed	8.00	
	Student	36.60	
	Unemployed	10.5	
	Prefer not to answer	6.80	
Income (PKR)	Below 25000	21.00	Average Monthly Household Income Rs 41545
	25,001 to 50,000	32.20	
	50,001 to 75,000	17.10	
	75,001 to 100,000	11.20	
	Above 100,000	18.50	

The validation of measurement and structural model for PLS path analysis was achieved employing Smart PLS 3.3.3. To assess the collinearity of single-source data, we measured the variance inflation factor (VIF) and found VIF between the acceptable range of five or less (VIF≤5), as directed by Hair et al. (2017).

5.2. Measurement model assessment

We applied internal consistency reliability, convergent validity, and discriminant validity criterion for the assessment of the measurement model (Hair et al., 2017). The criteria to determine the internal consistency reliability are labeled as Cronbach's alpha (α>0.60), Composite Reliability (CR>0.70), and Henseler's rho (ρA>0.70) (Ramayah et al., 2018), which are illustrated in Table 2.

Table 2. Measurement model assessment						
Construct	**Items**	**Loadings**	**CA**	**rho_A**	**CR**	**AVE**
PTS	PTS1	0.761	0.864	0.866	0.896	0.552
	PTS2	0.786				
	PTS3	0.716				
	PTS4	0.710				
	PTS5	0.794				
	PTS6	0.760				
	PTS7	0.666				
SMU	SMU1	0.637	0.800	0.811	0.856	0.500
	SMU2	0.690				
	SMU3	0.772				
	SMU4	0.785				
	SMU5	0.639				
	SMU6	0.706				
PEX	PEX1	0.774	0.869	0.878	0.897	0.522
	PEX2	0.813				
	PEX3	0.768				
	PEX4	0.697				
	PEX5	0.735				
	PEX6	0.687				
	PEX7	0.678				
	PEX8	0.607				
PI	PI1	0.856	0.848	0.850	0.892	0.625
	PI2	0.792				
	PI3	0.826				
	PI4	0.691				
	PI5	0.778				
PEM	PEM1	0.704	0.816	0.824	0.865	0.516
	PEM2	0.796				
	PEM3	0.672				
	PEM5	0.716				
	PEM6	0.698				
	PEM7	0.719				

Note: PEM (Political Empowerment), PEX (Political Expression), PI (Political Interest), PTS (Partisanship), SMU (Social Media Use), CA (Cronbach's Alpha), CR (Composite Reliability), AVE (Average Variance Extracted).

Outer loadings of the indicators and Average Variance Extracted (AVE) are recommended to evaluate the convergent validity of the construct. In this study, we kept items above 0.60 loading. We deleted only

those items necessary to meet the Average Variance Extracted (AVE>0.50) criterion or with the loading less than 0.5 (Hair et al., 2017). Following these steps, PEM4 with loading 0.629 was deleted to establish the AVE criterion of convergent validity for political empowerment. All the relevant values are listed in Table 2.

Discriminant validity ensures that each construct is unique and is empirically measuring a phenomenon different from other constructs in the same framework. We utilized the traditional Fornell and Larcker (1981) criterion and Heterotrait-Monotrait (HTMT) criterion to assess the discriminant validity. According to Fornell and Larcker, discriminant validity is established if the diagonal values of each construct are larger than its corresponding correlation coefficients. For the Heterotrait-Monotrait (HTMT) ratio of correlations, Henseler et al. (2015), in a discussion on the HTMT threshold level, stated that it is debatable. In this study, following Gold et al. (2001) and Teo et al. (2008) recommendations, we considered 0.90 (HTMT0.90) as the threshold level and met the discriminant validity (Table 3).

Table 3. Discriminant validity criteria

	Fornell-Larcker Criterion					HTMT Criterion				
	1	2	3	4	5	1	2	3	4	5
1.PEM	0.719									
2.PEX	0.376	0.722				0.429				
3.PI	0.562	0.393	0.791			0.649	0.450			
4.PTS	0.320	0.445	0.398	0.743		0.365	0.493	0.461		
5.SMU	0.546	0.359	0.591	0.339	0.707	0.657	0.410	0.720	0.403	

Note. PEM (Political Empowerment), PEX (Political Expression), PI (Political Interest), PTS (Partisanship), SMU (Social Media Use).

5.3. Structural model assessment

After fulfilling the requirements, we assessed the structural model to measure the model's predictive capability and relationship between constructs (Hair et al., 2017). To test the significance of the relationship, we followed Ramayah et al. (2018)'s directive and conducted bootstrapping of 410 cases with 5,000 samples. Results of direct and indirect relationships in structural path analysis, referring to the Pakistani young voters, are presented in Table 4. We found a positive and significant relationship between social media use and political expression (β=0.235, p=0.000, t=5.182). Hence, H1 secured substantial support. The path coefficient between partisanship and political expression (β=0.365, p=0.000, t=8.563) was positive and significant. The result is supportive towards H2 and suggests that respondents with partisan behavior have a tendency toward political expression. Similarly, political expression was found to be a positive predictor of political empowerment (β=0.160, p=0.000, t=3.727). Hence, H3 was also supported. In addition to β, p, and t values, the confidence interval bias-corrected (CIBC) results are also required to report the strength of path analysis (Rasoolimanesh et al., 2021). The significance of the relationship is determined if zero "0" does not fall between the lower and upper levels of the confidence interval (Ramayah et al., 2018). The results in Table 4 illustrate that H1, H2, and H3 met the CIBC criterion.

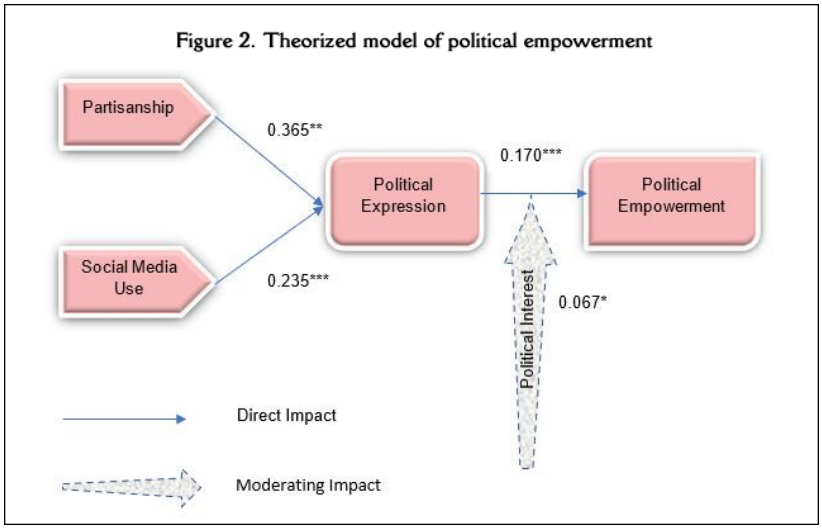
Table 4. Direct and indirect effect hypotheses

Relationships	β	t-value	p-value	CIBC LL=5%, UL=95%	Supported
SMU -> PEX	0.235	5.182	0.000***	[0.151, 0.303]	Yes
PTS -> PEX	0.365	8.563	0.000***	[0.289, 0.430]	Yes
PEX -> PEM	0.170	3.727	0.000***	[0.093, 0.241]	Yes
PEX*PI -> PEM	0.067	1.847	0.032*	[0.008, 0.126]	Yes

Note. β values, t-values, and p-values were computed through bootstrapping with 5000 samples, *p<0.05, **p<0.01, ***p<0.001.

We tested indirect effect relationships to determine the moderating role of political interest between political expression and political empowerment. Political interest proved as a significant moderator (β=0.067, p=0.032, t=1.847). Thus, H4 was also supported. The indirect hypotheses' results suggested that showing political interest as a moderator strengthens this relationship.

Furthermore, the findings of this study revealed that the framework's outcome variable has a moderate coefficient of determination (R^2). R^2 describes the percentage of variance in the dependent variable caused by the predicting variables. In this study, the R^2value for political empowerment was 0.345 (Figure 2), which depicts moderate and satisfactory predictive accuracy of the model (Hair et al., 2017).



6. Discussion

This article aimed at exploring the main research question of "whether political interest moderates/enhances the effect of the political expression on political empowerment of youngsters". The study also includes three other paths of relationships. First, it explores the direct relationship of social media with political expression. Second, it discovers the relationship between political partisanship and political expression. Third, it enquires about the effects of political expression on the political empowerment of youngsters. The study is structured based on the theoretical framework of rational choice theory from the work of George C. Homans (1958), Peter Blau (1964), and Tajfel's (1974) social Identity theory.

Multiple statistical procedures have been carried out to measure the direction, significance, and strength of the relationships. The findings identified a positive and significant path for the relationships of social media use with political expression, which is endorsed by many previous and contemporary studies (Bimber et al., 2015). The study findings also suggest a positive and significant relationship between partisanship of youngsters with their political expression. Several relevant studies provide support to these findings. Partisanship has proved a key variable that encourages political expression on various political issues (Dancey & Goren, 2010; Kim, 2018). These studies were carried out in the stable political environments and strong democracies of the United States and South Korea. However, the same outcome in the case of Pakistan's developing democracy in a turbulent political environment demonstrates the consistency of partisanship in influencing political expression. Moreover, the results show that political expression in any form increases the likelihood of political empowerment of youngsters.

The goal of this study was to explore the twofold requirements of an overarching theoretical framework. The study findings are sufficiently in accordance with rational choice theory, which poses high reliance on deliberate choices of individuals in selection matters. Findings suggest political interest as a substantial moderator that enhances the effect of political expression on the political empowerment of youngsters. Meanwhile, the theory about partisanship i.e., SIT, recalls the behavioral position of an individual as a cohort. They are more likely to express themselves in an in-group setting having a like-minded ideology. Political partisanship tends to exhibit the same political norms in political groups of the same ideology which SIT suggests.

In Pakistan, there is a widespread belief that young, educated people, particularly females, are disinterested in politics. There are multiple explanations for this political inactivity. The absence of political training in the form of student unions, weak economic conditions, political victimization by institutions, and intensive screen use are some of the reasons consistently presented (Hassan & Sabir, 2020). In contrast, the findings of this study negated all these. Most of the participants in this research were graduates from universities, 47.80% were female, and almost all were frequent social media users. Considering all possible reasons for political deactivation, as mentioned earlier, the findings of this study revealed that

the political interest of respondents develops a sense of political empowerment among social media users. From the above discussion, we conclude that if the aim is to establish a democratic political structure in the country and the democratic strength of citizens for the stability of the political setup, political interest is a very substantial element in catalyzing youngsters' political empowerment. Although the role of political interest in shaping political behaviors has been studied in the past, its moderating role in enhancing political empowerment and moreover, in the context of Pakistan was novel in the literature.

Based on the current research, that primarily highlighted the importance of social media in democracy and its effect on political empowerment, the importance of political expression for political empowerment is also determined. This study suggests that policymakers in Pakistan should encourage social media and political forums so that citizens can express themselves as much as possible.

7. Limitations and recommendations

The main limitation of this study is related to data collection challenges caused by the COVID-19 pandemic. Initially, the multistage cluster sampling approach was intended for data collection. This technique required physical administration which was not viable at that time. Overcoming this hurdle, we applied snowball sampling, which resulted in a change in the predetermined data collection method. Second, this study is cross-sectional. If data is collected in a longitudinal research design during and after elections, the findings may vary. Data collection during the election campaign will help elucidate the effects of the negative and positive political campaigns.

Referring to the demographic distribution of the Pakistani population, 62.60% population of the country lives in rural areas and 37.40% in urban areas (World Bank, 2021). Therefore, this study recommends future researchers to conduct a multi-group analysis (MGA) to compare the political characteristics of the rural and urban population of Pakistan.

Authors' Contribution

Idea, T.R, Z.F; Literature review (state of the art), Z.F; Theoretical Framework, Z.F; Methodology, T.R; Data analysis, T.R; Results, T.R; Discussion and conclusions, T.R, Z.F; Writing (original draft), T.R, Z.F; Final revisions, T.R., Z.F; Project design, T.R., Z.F.

References

Ahmad, A., Bhatti, M.I., & Yousaf, F.F. (2020). Whom to vote? Socio-psychological factors influencing voting behavior in Rural Punjab. *Pakistan*. *Sir Syed Journal of Education & Social Research*, *3*(2), 9-15. https://doi.org/10.36902/sjesr-vol3-iss2-2020(9-15)

Ali, F., Rasoolimanesh, S.M., Sarstedt, M., Ringle, C.M., & Ryu, K. (2018). An assessment of the use of partial least squares structural equation modeling (PLS-SEM) in hospitality research. *International Journal of Contemporary Hospitality Management*, *30*(1), 514-538. https://doi.org/10.1108/IJCHM-10-2016-0568

Ali, R., & Fatima, A. (2016). Social media, newspapers and political awareness of youth in Pakistan. *Pakistan Library & Information Science Journal*, *47*(1), 30-41. https://bit.ly/3bkxuHY

Ashraf, D., Ali, T., & Hosain, A. (2013). Youth development and education in Pakistan: Exploring the relationship. *Sisyphus: Journal of Education*, *1*(2), 162-192. https://bit.ly/3cCT76I

Barnhart, B. (2021). *Social media demographics to inform your brand's strategy in 2021*. Sprout Social. https://bit.ly/3S6i7DD

Becker, A.B., & Copeland, L. (2016). Networked publics: How connective social media use facilitates political consumerism among LGBT Americans. *Journal of Information Technology & Politics*, *13*(1), 22-36. https://doi.org/10.1080/19331681.2015.1131655

Bimber, B., Cunill, M.C., Copeland, L., & Gibson, R. (2015). Digital media and political participation: The moderating role of political interest across acts and over time. *Social Science Computer Review*, *33*(1), 21-42. https://doi.org/10.1177/0894439314526559

Blau, P.M. (1964). Justice in social exchange. *Sociological Inquiry*, *34*(2), 193-206. https://doi.org/10.1111/j.1475-682X.1964.tb00583.x

Bobbie, E. (2013). *The practice of social research*. Pearson Education Limited.

Boulianne, S. (2015). Social media use and participation: A meta-analysis of current research. *Information, Communication and Society*, *18*, 524-538. https://doi.org/10.1080/1369118X.2015.1008542

Chan, M. (2018). Partisan strength and social media use among voters during the 2016 Hong Kong legislative council election: Examining the roles of ambivalence and disagreement. *Journalism & Mass Communication Quarterly*, *95*(2), 343-362. https://doi.org/10.1177/1077699017750857

Chen, Z., & Chan, M. (2017). Motivations for social media use and impact on political participation in China: A cognitive and communication mediation approach. *Cyberpsychology, Behavior, and Social Networking*, *20*(2), 83-90. https://doi.org/10.1089/cyber.2016.0174

Colás-Bravo, P., González-Ramírez, T., & De-Pablos-Pons, J. (2013). Young people and social networks: Motivations and preferred uses. [Juventud y redes sociales: Motivaciones y usos preferentes]. *Comunicar*, *40*, 15-23. https://doi.org/10.3916/C40-2013-02-01

Dancey, L., & Goren, P. (2010). Party identification, issue attitudes, and the dynamics of political debate. *American Journal of Political Science*, *54*(3), 686-699. https://doi.org/10.1111/j.1540-5907.2010.00454.x

Foos, F., & De-Rooij, E.A. (2017). The role of partisan cues in voter mobilization campaigns: Evidence from a randomized field experiment. *Electoral Studies*, *45*, 63-74. https://doi.org/10.1016/j.electstud.2016.11.010

Fornell, C., & Larcker, D.F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39-50. https://doi.org/10.1177/002224378101800104

Gil-De-Zúñiga, H., Copeland, L., & Bimber, B. (2014). Political consumerism: Civic engagement and the social media connection. *New Media and Society*, *16*(3), 488-506. https://doi.org/10.1177/1461444813487960

Gil-De-Zúñiga, H., Molyneux, L., & Zheng, P. (2014). Social media, political expression, and political participation: Panel analysis of lagged and concurrent relationships. *Journal of Communication*, *64*(4), 612-634. https://doi.org/10.1111/jcom.12103

Gold, A.H., Malhotra, A., & Segars, A.H. (2001). Knowledge management: An organizational capabilities perspective. *Journal of Management Information Systems*, *18*(1), 185-214. https://doi.org/10.1080/07421222.2001.11045669

Großer, J., & Schram, A. (2006). Neighborhood information exchange and voter participation: An experimental study. *American Political Science Review*, *100*(2), 235-248. https://doi.org/10.1017/S0003055406062137

Hair, J.F., Hault, G.T.M., Ringle, C.M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)*. SAGE Publications Inc. https://bit.ly/3Saj4Lc

Hair, J.F., Ringle, C.M., Gudergan, S.P., Fischer, A., Nitzl, C., & Menictas, C. (2019). Partial least squares structural equation modeling-based discrete choice modeling: An illustration in modeling retailer choice. *Business Research*, *12*(1), 115-142. https://doi.org/10.1007/s40685-018-0072-4

Harrop, M., & Miller, W.L. (1987). *Elections and voters: A comparative introduction*. Meredith Press. https://doi.org/10.1007/978-1-349-18912-0

Hassan, T., & Sabir, M. (2020). Civil society-democracy nexus in Pakistan. *South Asian Studies*, *26*(2), 325-347. https://bit.ly/3bfRXhg

Henseler, J., Ringle, C.M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115-135. https://doi.org/10.1007/s11747-014-0403-8

Himelboim, I., Lariscy, R.W., Tinkham, S.F., & Sweetser, K.D. (2012). Social media and online political communication: The role of interpersonal informational trust and openness. *Journal of Broadcasting & Electronic Media*, *56*(1), 92-115. https://doi.org/10.1080/08838151.2011.648682

Homans, G.C. (1958). Social behavior as exchange. *American Journal of Sociology*, *63*(6), 597-606. https://doi.org/10.1086/222355

Huddy, L., Mason, L., & Aaroe, L. (2010). Measuring partisanship as a social identity, predicting political activism. In *Annual Meeting of the International Society for Political Psychology* (pp. 1-33). ISPP. https://bit.ly/3PN71Sv

Kalogeropoulos, A., Negredo, S., & Picone, I. (2017). Who shares and comments on news?: A cross-national comparative analysis of online and social media participation. *Social Media + Society*, *3*(4). https://doi.org/10.1177/2056305117735754

Kim, M. (2018). How does Facebook news use lead to actions in South Korea? The role of Facebook discussion network heterogeneity, political interest, and conflict avoidance in predicting political participation. *Telematics and Informatics*, *35*(5), 1373-1381. https://doi.org/10.1016/j.tele.2018.03.007

Krejcie, R.V., & Morgan, D.W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement*, *30*(3), 607-610. https://doi.org/10.1177/001316447003000308

Loader, B.D., Vromen, A., & Xenos, M.A. (2014). The networked young citizen: social media, political participation and civic engagement. *Information, Communication and Society*, *17*, 143-150. https://doi.org/10.1080/1369118X.2013.871571

Maurissen, L. (2020). Political efficacy and interest as mediators of expected political participation among Belgian adolescents. *Applied Developmental Science*, *24*(4), 339-353. https://doi.org/10.1080/10888691.2018.1507744

Moffett, K.W., & Rice, L.L. (2018). College Students and online political expression during the 2016 election. *Social Science Computer Review*, *36*(4), 422-439. https://doi.org/10.1177/0894439317721186

Morrell, M.E. (2005). Deliberation, democratic decision-making and internal political efficacy. *Political Behavior*, *27*, 49-69. https://doi.org/10.1007/s11109-005-3076-7

Pakistan Bureau of Statistics (Ed.) (2021). *Demographic and research*. https://bit.ly/3viHvwj

Prior, M. (2019). *Hooked: How politics captures people's interest*. Cambridge University Press. https://doi.org/10.1017/9781108355001

Ramayah, T., Cheah, J., Chuah, F., Ting, H., & Memon, M.A. (2018). *Partial Least squares structural equation modeling (PLS-SEM) Using SmartPLS 3.0 (2nd (ed.)*. Pearson.

Rasoolimanesh, S.M., Khoo-Lattimore, C., Md-Noor, S., Jaafar, M., & Konar, R. (2021). Tourist engagement and loyalty: Gender matters? *Current Issues in Tourism*, *24*(6), 871-885. https://doi.org/10.1080/13683500.2020.1765321

Riaz, S. (2007). Government advertisements - Influence on print media content: A content analysis of the leading newspapers of Pakistan. *Asian Journal of Social Science*, *35*(2), 154-178. https://doi.org/10.1163/156853107X203414

Sarstedt, M., Hair, J.F., Cheah, J.H., Becker, J.M., & Ringle, C.M. (2019). How to specify, estimate, and validate higher-order constructs in PLS-SEM. *Australasian Marketing Journal*, *27*(3), 197-211. https://doi.org/10.1016/j.ausmj.2019.05.003

Shively, W.P. (1979). The relationship between age and party identification: A cohort analysis. *Political Methodology*, *6*(4), 437-446. https://bit.ly/3olJRAO

Skoric, M.M., Zhu, Q., Goh, D., & Pang, N. (2016). Social media and citizen engagement: A meta-analytic review. *New Media and Society*, *18*(9), 1817-1839. https://doi.org/10.1177/1461444815616221

Spreitzer, G.M. (1995). Psychological empowerment in the workplace: Dimensions, measurement, and validation. *Academy of Management Journal*, *38*, 1442-1465. https://doi.org/10.5465/256865

Tajfel, H. (1974). Social identity and intergroup behaviour. *Information (International Social Science Council)*, *13*(2), 65-93. https://doi.org/10.1177/053901847401300204

Tariq, R., Zolkepli, I.A., & Ahmad, M. (2022). Political participation of young voters: Tracing direct and indirect effects of social media and political orientations. *Social Sciences*, (2), 11-11. https://doi.org/10.3390/socsci11020081

Teo, T.S.H., Srivastava, S.C., & Jiang, L. (2008). Trust and electronic government success: an empirical study. *Journal of Management Information Systems*, *25*(3), 99-132. https://doi.org/10.2753/MIS0742-1222250303

United Nations Development Pakistan (Ed.) (2018). *National Human Development Report 2017: Unleashing the potential of a young Pakistan*. UNESCO. https://bit.ly/3zNiDiI

Van-Dop, N., Depauw, J., & Driessens, K. (2016). Measuring empowerment: Development and validation of the service user psychological empowerment scale. *Journal of Social Service Research*, *42*(5), 651-664. https://doi.org/10.1080/01488376.2016.1216915

World Bank (Ed.) (2021). *World development indicators*. DataBank. https://bit.ly/3BmsXQ7

Zeib, F. (2022). Rising wave of social media: Empowering university students through online and offline political participation. *Webology*, *19*(2), 4540-4558. https://bit.ly/3BmzL07

Zolkepli, I.A., Kamarulzaman, Y., & Kitchen, P.J. (2018). Uncovering Psychological gratifications affecting social media utilization: A multiblock hierarchical analysis. *Journal of Marketing Theory and Practice*, *26*(4), 412-430. https://doi.org/10.1080/10696679.2018.1489730

Zulqarnain, W., & Taimur-Ul-Hassan (2017). Social and traditional media usage: A demographic analysis of Pakistani youth. *Global Media Journal: Pakistan Edition*, *10*(1), 30-56. https://bit.ly/3OGLcCP

Comunicar, n. 74, v. XXXI, 2023 | Media Education Research Journal | ISSN: 1134-3478; e-ISSN: 1988-3478
www.comunicarjournal.com

Digital competence among young people in Spain: A gender divide analysis

Competencias digitales de la juventud en España: Un análisis de la brecha de género

Dr. Elisenda Estanyol. Associate Professor, Faculty of Information and Communication Sciences, Universitat Oberta de Catalunya, Barcelona (Spain) (eestanyol@uoc.edu) (https://orcid.org/0000-0003-3986-0377)
Dr. Mireia Montaña. Associate Professor, Faculty of Information and Communication Sciences, Universitat Oberta de Catalunya, Barcelona (Spain) (mmontanabl@uoc.edu) (https://orcid.org/%200000-0002-6464-5129)
Pedro Fernández-de-Castro. Doctoral Candidate, Programme in Humanities and Communication, Universitat Oberta de Catalunya, Barcelona (Spain) (pfernandez_de_castro@uoc.edu) (https://orcid.org/0000-0003-3994-9953)
Dr. Daniel Aranda. Associate Professor, Faculty of Information and Communication Sciences, Universitat Oberta de Catalunya, Barcelona (Spain) (darandaj@uoc.edu) (https://orcid.org/0000-0001-9290-0708)
Dr. Leila Mohammadi. Collaborating Researcher, GAME Research Group, Universitat Oberta de Catalunya, Barcelona (Spain) (limohammadi@uoc.edu) (https://orcid.org/0000-0001-9090-2427)

ABSTRACT

Datafication in today's communicative ecosystem poses a challenge to media and digital literacy, especially with regard to young people's participation and civic and democratic engagement. We address this issue using the notion of digital citizenship, in order to study the gender digital divide as it relates to competence (i.e., skills and knowledge) and the possibility of leveraging said competence to promote civic education grounded in gender equality in the digital environment. For this study, we surveyed a representative sample of 600 young people between the ages of 16 and 18 in Spain to gauge their digital competence through three variables: technical skills, informational skills and critical knowledge. We then performed a descriptive analysis resulting in percentages, means and standard deviations and a bivariate analysis using significance testing (T-tests) between the above variables and the gender variable. Results show a relative balance between men and women in technical and informational digital skills, albeit tipped slightly in favour of women. By contrast, men claim to have more critical knowledge. Based on these results, we discuss the need to consider the contributions of feminist theories in the field of technology to develop proposals for teaching digital competence that encourage active digital citizenship based on gender equality.

RESUMEN

El escenario de datificación del ecosistema comunicativo actual plantea un desafío a la alfabetización mediática y digital, especialmente en lo que respecta a la participación y el compromiso cívico y democrático de la población joven. En este artículo abordamos esta cuestión a partir de la noción de ciudadanía digital con el objetivo de estudiar la brecha digital de género en términos de capacidades —competencias y conocimientos— y la posibilidad de aprovecharlas para promover una educación cívica fundamentada en la igualdad de género en el entorno digital. Para ello, mediante una encuesta con una muestra representativa de 600 personas jóvenes —entre 16 y 18 años— en España, observamos su nivel de competencias digitales a través de tres variables: competencias técnicas, competencias informacionales y conocimientos críticos. Los resultados del análisis descriptivo, mediante porcentajes, medias y desviaciones típicas; y bivariado entre dichas variables y la variable de género mediante pruebas T-test de significatividad, muestran que, si bien hay relativa igualdad de condiciones entre hombres y mujeres en cuanto a competencias digitales técnicas e informacionales con un ligero dominio de las mujeres, los hombres afirman tener más conocimientos críticos. Desde ahí, discutimos la necesidad de considerar los aportes de las teorías feministas en el ámbito tecnológico para elaborar propuestas educativas en competencias digitales que fomenten desde la igualdad de género la promoción de una ciudadanía digital activa.

KEYWORDS | PALABRAS CLAVE

Digital citizenship, citizenship education, media literacy, critical thinking, gender equality, young people.
Ciudadanía digital, educación ciudadana, competencia mediática, pensamiento crítico, igualdad de género, jóvenes.

Received: 2022-05-30 | Reviewed: 2022-06-30 | Accepted: 2022-07-13 | OnlineFirst: 2022-10-01 | Published: 2023-01-01
DOI https://doi.org/10.3916/C74-2023-09 | Pages: 107-116

1. Introduction and state of the art

Digital platforms are taking over as the predominant venues for shaping public opinion, effectively transforming the media ecosystem and posing significant challenges for media literacy (Valtonen et al., 2019). These corporate platforms employ a business model based on personal data collection and traffic, leaning into the paradigm of datafication, i.e., the transformation of social actions into quantifiable online data, using methods that implement "dataveillance", which involves surveilling people through the use of their online data (Van-Dijck, 2014). When the Cambridge Analytica case came to light in 2018, it became known that the company had used personal data to build psychometric profiles to influence electoral processes such as Brexit and the 2016 United States presidential election. This scandal came as a wake-up call about the impact of big data and digital-platform-managing algorithms on democratic societies. Within the discourse on the true effect of algorithm-driven "echo chambers" (Dubois & Blank, 2018), we subscribe to the idea that these shape "immersive media environments" that affect people's values and actions (Cohen, 2018).

Debates on what is meant by digital citizenship and how it can be promoted through education date back to the beginning of the 21st century (Richardson et al., 2021), but they are now more relevant than ever given the current state of affairs. This article is a response to the need to digitally empower people to face the media challenges that question democracy as a system of social organisation. Our research focuses on young people aged 16 to 18 in their position as citizens close to acquiring full participation rights. We also specifically hone in on the gender factor, in order to explore whether digital technologies contribute evenly to young people's civic engagement.

The research questions guiding our study are as follows: [RQ1] How digitally competent do young Spaniards claim to be? [RQ2] Are there gender differences in young people's self-perceived digital competence? Our hypotheses are: [H1] Young Spaniards will report having better technical and informational digital skills than critical digital knowledge. [H2] There will be differences between young men and women regarding their self-perceived digital competence, especially with respect to critical knowledge.

1.1. Digital citizenship and youth

Digital citizenship can be understood as a multidimensional concept encompassing literacy, competencies, participation, and digital access and divides. However, it is still being defined and addressed in various fields, including education, communication and political science (Gleason & Von-Gillern, 2018; Panke & Stephens, 2018). From this discourse, we agree with Richardson et al.'s (2021) systematic review of digital citizenship, which found the most comprehensive paper to be by Choi (2016). This author defines digital citizenship as "abilities, thinking and action regarding internet use" and says that it "allows people to understand, navigate, engage in, and transform self, community, society, and the world" (Choi, 2016: 20).

Relevant here is the proposal by Yue et al. (2019), which identifies two predominant approaches to digital citizenship in relation to youth. First is the "control approach", which conceives young people as passive subjects and "not-yet-citizens", with digital citizenship adopting a normative perspective. This approach excludes young people from participation mechanisms and subordinates them to adult-centric criteria. Second is the "freedom approach", which views young people as active agents, emphasising participation and inclusion through digital media. However, this perspective assumes that young people are digital natives and autonomous users in open and horizontal participatory processes, which implies understanding digital participation as inherently positive and taking an uncritical view of the datafication scenario presented above. Again, drawing on Yue et al. (2019), we embrace a third approach that emphasises the civic aspect of digital participation and the importance of media literacy. This perspective requires a critical understanding of young people, not as passive subjects or active agents per se, but as users embedded in a complex, technical, commercial, cultural, social, and political assemblage in two interrelated environments (offline and online). In practical terms, this approach calls for us to examine the youth population's digital practices and, most crucially, any differences within this demographic (Darvin, 2018; Porat et al., 2018).

1.2. The digital divide and digital competence from a gender perspective

Taking this approach, we draw on Choi's (2016) conceptual analysis of digital citizenship, which identifies four major categories: "ethics", "media and information literacy", "participation/engagement" and "critical resistance". This article expands on the "media and information literacy" category by exploring the three identified sub-themes: the digital divide, technical skills and psychological capabilities.

The digital divide concept emerged when digital technologies became available and widespread. Their social impact soon became apparent, namely through an uneven distribution among the population, whether by gender, class, ethnicity, location, level of education, or other socio-economic and socio-cultural factors. This phenomenon is also more explicitly referred to under the term "digital inequalities", which highlights the link between digital and social exclusion (Helsper, 2021). Nomenclature aside, research in this area has evolved from technical enquiries about who has access to digital devices and who does not, to multidimensional studies considering differences in digital uses and skills as a "second-level digital divide" (Hargittai, 2002). Recently, research on socio-digital inequalities has explored how "digital capital" (Ragnedda, 2016), digital resources, (internet connection, devices, etc.) and digital capabilities (skills, competencies and knowledge) can lead to "tangible" benefits outside the digital environment, be they personal, political, cultural, financial, or social, in what amounts to a "third-level digital divide" (Van-Deursen & Helsper, 2015; Ragnedda, 2016). On the gender factor in the digital divide, it needs to be said that technology has historically been a predominantly male domain (Cockburn, 1983; Wajcman, 2004; Loh, 2019). Research on the digital gender divide is therefore a continuation of these studies (Gurung, 2018).

Thanks to this multi-level, digital competence-embracing perspective, we are able to link the digital divide to the other two media and information literacy sub-themes outlined by Choi (2016). First, technical skills, which in this article we refer to as "technical digital skills". According to Choi (2016), these "represent an instrumental perspective on literacies and competencies, such as how to use new digital technologies, computers, smartphones and/or tablet PCs" and "serve as prerequisites for advanced internet activities" (14). Second, psychological capabilities, referred to in this article as "informational digital skills". These encompass "cognitive-intellectual abilities to select, classify, analyse, interpret, and understand data critically [...]; socio-communicative abilities to communicate/network with others, share photos/videos, or exchange ideas through blogs, podcasts, and/or online discussion forums [...]; and emotional abilities to learn how to control negative feelings or sympathize with others' emotions" (Choi, 2016: 14-15). Both types of skills can be interpreted at the second level of the digital divide.

While Choi's (2016) proposal does state that digital competencies cannot be reduced to their instrumental aspect, a consideration shared by Gutiérrez-Martín and Tyner (2012), we also agree with the latter authors in that media education cannot be reduced to digital competencies either (ibid). We therefore extend our analysis to include a third component relating to "critical digital knowledge". This is based on the premise that, although media literacy and digital literacy belong to the same field, their key difference lies on the particular focus that the digital aspect places on the structural transformations imposed on society by the digital environment (Bali, 2019). For this purpose, we follow Mihailidis et al.'s (2021) proposal to take "critical consciousness" as the central value of a transformative media pedagogy (here in its digital form), adopting the Freirean approach of "conscientisation" (Freire, 1970), understood as the "liberating education process through which people [...] acquire a critical awareness of themselves and reality, which they turn into action, thus affirming themselves as conscious subjects and co-creators of their historical future" (Díez-Gutiérrez, 2022: 51). Based on this perspective, we add a scale of digital environment-related knowledge that is considered critical, insofar as it allows users to distance themselves enough to structurally understand how digital media work.

By further developing this pedagogical approach reinterpreted by Hooks (1994) from an intersectional feminist perspective, we are able to lay the foundations for a transgressive educational proposal. We discuss the need and possible features of a type of literacy that draws from the media tradition and focuses on embracing a critical (Pangrazio, 2016) and feminist perspective (Bali, 2019) when engaging with the digital environment. A type of literacy that is not only geared towards employment in the labour market, but also the formation of digital citizenship (Pötzsch, 2019).

2. Material and methods

2.1. Design

Research design is quantitative and cross-sectional, involving an online survey intended to measure the self-perceived digital competence of young people aged 16 to 18 living in Spain. This age group was chosen because they are close to acquiring full citizen participation rights, such as the right to vote in Spain. The survey was self-administered, i.e., it was completed by the respondents themselves without the presence of an interviewer, between 23 September and 5 October 2021, with prior informed consent. Data and participant security and confidentiality were respected following the UNE EN ISO/IEC 27001 standards and the favourable report issued by the Universitat Oberta de Catalunya (UOC) Ethics Committee under file CE22-PR05.

2.2. Survey

The survey is based on Van-Deursen et al.'s (2016) and Aranda et al.'s (2020) extension of the Oxford Internet Institute's (OxIS) WIP Britain 2013, combined with a systematic review of the notion of "digital youth work" (Fernández-de-Castro et al., 2021). It includes a section of socio-demographic questions (gender, age, place of residence, etc.) and three further sections comprising 24 questions about the respondents' self-perceived digital competence. Because the survey asked young people about how they perceive their own digital skills and knowledge, their answers may not necessarily match their actual competence level.

The first and second sections on digital competence asked about technical skills (nine items) and informational skills (ten items), respectively. Answers were given on a 5-point Likert scale: 1: I don't know what this is or what it means; 2: I know what this is but I don't know how to do it; 3: I would know how to do this with help; 4: I know how to do this by myself; and 5: I know how to do this and could teach others. The third section dealt with critical knowledge of the digital environment and included five items, also measured on a 5-point Likert scale: 1: Nothing at all; 2: A little bit; 3: An average amount; 4: A fair amount; and 5: A lot.

A principal component analysis was carried out on the proposed scales to check their validity and Cronbach's alpha was used to measure their reliability. Regarding the section on technical digital skills, the analysis showed an acceptable structure for all nine items (KMO=0.910; Bartlett's test significant with $p<0.001$). The structure comprised two components explaining 64.8% of the total variance (40.1% for the first component and 24.7% for the second); the Cronbach's alpha coefficient was 0.903 for the first component and 0.773 for the second. For the section on informational digital skills, the analysis showed an acceptable structure for all ten items (KMO=0.955; Bartlett's test significant with $p<0.001$). A single component explained 59.9% of the total variance and Cronbach's alpha was 0.925. For the section on critical digital knowledge, the analysis showed an acceptable structure for all five items (KMO=0.843; Bartlett's test significant with $p<0.001$). A single component explained 58% of the total variance and Cronbach's alpha was 0.819.

2.3. Sample

Data collection was carried out by ODEC, which sent a link to the participants, fulfilling representative criteria in terms of age, gender, education, and place of residence in Spain. By means of a simple random sampling strategy, 600 young people completed the survey, taking an average of 13 minutes per person. The sampling response rate was 62.11%, with a margin of error of 4% for the sample as a whole, a confidence level of 95% (1.96 standard deviation) and maximum indeterminacy (P=Q=50%). Subsequently, the stratification was weighted to refine the respondent weights based on the population data of the final study universe. The reference data for the weighting coefficient were calculated using the variables "Nielsen area", "municipality size", "gender", and "age" from the latest wave of the Spanish General Media Study (EGM).

3. Analysis and results

The results were processed using IBM SPSS Statistics 24®. We first performed a descriptive statistical analysis of the survey's Likert scale variables, including calculating means and standard deviations. We also carried out a bivariate analysis to check the digital competence variables against the gender variable. For inference, we performed an independent samples T-test to determine whether there was any significant difference between two groups, with the significance level set at 5%. In what follows, we highlight key data relating to the respondents' autonomously self-reported skills. We also cover the knowledge items, focusing on the figures for knowing a lot or a fair amount. The remaining data are presented in the tables below.[1]

With regard to technical skills (Table 1), the results show that 73.4% of young people claim to know how to install/uninstall basic programs and applications without help. There is hardly any difference by gender in this regard (73.3% of women; 73.5% of men). Most young people report knowing how to browse the internet and use related services for everyday purposes, with 80.7% saying they can do this without help. Again, there is no significant difference by gender (80.7% of women; 80.9% of men).

Fewer young people appear to use content management platforms to produce multimedia publications, with only 26.6% claiming to know how to do this without help. We do not observe a relevant difference by gender in this respect. Meanwhile, 58.4% of the respondents say they know how to record, edit and upload video content to the internet without help. In terms of difference by gender, young women claim to be more competent in this regard than young men (60.6% vs 56.2%). As for sharing and distributing digital multimedia content, 69.7% of the young people surveyed say they can do this without help, with no relevant difference by gender (69.5% of women; 69.9% of men). A total of 66.6% of young people say they know how to work with others using digital collaboration tools without help; by gender, this breaks down to 65% of young women and 68.2% of young men. Only 27.8% of young people say they know how to set up digital services and use tools to increase online privacy and anonymity without needing help; by gender, this is 29.3% of young men compared to 26.3% of young women. In terms of knowing how to read and/or write computer code, only 19.2% say they can do this without help, with no relevant gender difference (19% of women; 19.4% of men). Similarly, few young people claim to know how to repair and/or service devices without help (25.5%). There is a gender difference here, with the scales tipped in favour of young men (32.6%) compared to young women (18.3%).

Table 1. Young people's technical skills by gender (%)

Technical skills / gender	Women			Men			Total		
	I know how to do this	I would know how to do this with help	I don't know how to do this	I know how to do this	I would know how to do this with help	I don't know how to do this	I know how to do this	I would know how to do this with help	I don't know how to do this
Install/Uninstall basic programs and applications for my needs	73.3	12.1	6.9	73.5	10.2	9.9	73.4	11.1	8.5
Browse the internet and use related services for everyday purposes	80.7	8.5	7	80.9	6.7	7	80.7	7.6	6.9
Use content management platforms to produce multimedia publications	26.9	16.8	31.5	26.4	25.7	31.4	26.6	21.3	31.5
Record, edit and upload video content to digital platforms	60.6	17.8	12.2	56.2	24.6	12	58.4	21.3	12.1
Share and distribute multimedia content across social media, platforms and mailing lists	69.5	12.8	8.4	69.9	9.3	13.8	69.7	11	11.2
Work with others using digital collaboration tools	65.0	14.7	11	68.2	14.4	11.9	66.6	14.5	11.4
Set up digital services and use tools to increase online privacy and anonymity	26.3	24.8	31.8	29.3	24.8	30.2	27.8	24.8	31
Read and/or write computer code	19	19.6	37.4	19.4	27.8	32.9	19.2	23.8	35.1
Repair and/or service devices	18.3	21.7	30.1	32.6	32	16.9	25.5	26.8	23.4

In terms of informational skills (Table 2), 49.4% of young people say they know how to check the reliability and truthfulness of information without help. Broken down by gender, 48.9% of young women

claim to have this skill, one percentage point lower than young men. Among young people, 54% say they know how to classify and filter information to suit their interests without help. By gender, more women report having this skill than men (55.1% vs 53%). A total of 69.9% of young people say they are able to find and save information for use when they need it. A considerable difference appears when this figure is broken down by gender: 73.3% of young women claim to have this skill compared to 66.7% of young men. With respect to social informational skills, 64.4% of the respondents say they know how to display self-control when interacting with others on social media and digital forums so as not to react impulsively. Regarding spotting so-called "trolls" in online discussions, 58.5% of the young people surveyed claim to have this skill (54.9% of women; 61.9% of men), while 47.3% report knowing how to tell when they are interacting with a bot (41.9% of women; 52.5% of men).

Table 2. Young people's informational skills by gender (%)

Informational skills / gender	Women			Men			Total		
	I know how to do this	I would know how to do this with help	I don't know how to do this	I know how to do this	I would know how to do this with help	I don't know how to do this	I know how to do this	I would know how to do this with help	I don't know how to do this
Check the reliability and truthfulness of information	48.9	21.6	16.3	49.9	23.5	18.8	49.4	22.6	17.5
Classify and filter information	55.1	19.7	15.5	53.0	17.6	21.3	54	18.6	18.4
Find and save information	73.3	9.6	8.5	66.7	14.6	14.4	69.9	12.2	11.5
Display self-control when interacting with others on social media and digital forums	64.4	9.8	14.7	64.3	13.9	15.4	64.4	11.9	15.1
Identify users who act in a provocative manner ("trolls")	54.9	12.7	17.3	61.9	15.2	16	58.5	14	16.7
Know when I am interacting with a bot	41.9	15.8	25.7	52.5	16.9	19.4	47.3	16.3	22.5
Manage the various profiles that make up my digital identity	71.2	8.7	11	66.4	17.3	11.8	68.8	13	11.5
Adapt my behaviour to the standards of each platform	67.7	9.3	12	68.7	11.1	12.1	68.2	10.2	12.1
Identify my needs and find tools and platforms to fulfil them	59.1	12.5	16.1	54.5	19.4	16.4	56.8	16	16.2
Participate in online deliberation and decision-making processes	49.0	17.8	17.7	42.1	25.7	19.5	45.5	21.8	18.7

According to this sample, 68.8% of young people (71.2% of women; 66.4% of men) are able to manage the various profiles that make up their digital identity. Meanwhile, 68.2% say they know how to adapt their behaviour according to the standards of each platform. By gender, men come in one point above women in this respect. Among the young people surveyed, 56.8% report being able to identify their needs and find tools and platforms to fulfil them without help; women claim to be more competent in this skill than men (59.1% vs 54.5%). Less than half of the young people in our sample (45.5%) say they are able to take part in online deliberation and decision-making processes; 17.6% say they know how to do this and could teach others, while 27.9% say they simply know how to do this alone. By gender, women again come out ahead (49.0% vs 42.1%).

In terms of critical knowledge (Table 3), 22.9% of the respondents say they know a lot or a fair amount about the basic features of digital services; men stand out considerably when this figure is broken down by gender (27.1% vs 18.5%). Of the young people surveyed, 33.6% say they know a lot or a fair amount about how technology companies use personal data, with men claiming to know more than women (38.4% vs 28.6%). Meanwhile, 18.7% of young people say they know a lot or a fair amount about laws dealing with issues related to digital technologies.

By gender, the percentage of men is higher in this respect, at 24.6% compared to 18.8% for women. Only 22.4% say they know a lot or a fair amount about the influence of technology companies on public policy, a figure that breaks down to 18.7% of women and 25.9% of men. Finally, 31.9% of young people say they know how the technological devices they use are manufactured; by gender, more men self-reported such knowledge than women (36.9% vs 26.8%).

Table 3. Young people's digital knowledge by gender (%)

Digital knowledge / gender	Women			Men			Total		
	A lot or a fair amount	An average amount	A little or nothing	A lot or a fair amount	An average amount	A little or nothing	A lot or a fair amount	An average amount	A little or nothing
The basic features of digital services	18.5	23.6	46	27.1	29.4	36.8	22.9	26.6	41.4
The use of personal data by technology companies	28.6	30.3	33.3	38.4	28.4	28.1	33.6	29.3	30.7
Laws dealing with internet-related topics and digital technologies	18.8	27.4	42.1	24.6	29.4	41	21.7	28.4	41.5
The influence of technology companies on policy	18.7	27.2	41.7	25.9	26.1	38.9	22.4	26.7	40.2
How technological devices are manufactured	26.8	29.8	30.4	36.9	31.2	27.1	31.9	30.5	28.6

3.1. Bivariate analysis

The results of the bivariate analysis[2] (Table 4) reveal trends indicating competence differences between men and women in two of the three dimensions. The technical skills dimension shows similar results by gender; men are self-reportedly more competent in five of the items while women, in the remaining four. However, the difference is more pronounced in the items favouring men, ranging from 11 to 44 percentage points. The largest gap is in the item on repairing and/or servicing devices. Meanwhile, the four items leaning in women's favour show a slighter difference of between 2 and 13 percentage points.

In terms of informational skills, women self-reported higher competence in almost all the items, with differences of between 7 and 26 percentage points in the mean scores. In addition, the results of the statistical T-test show significant differences by gender in two items in this dimension, namely in the ability to "find and save information for use when needed" (P=0.026, N=600) and to "display self-control when interacting with others on social media and digital forums so as not to react impulsively" (P=0.009, N=600).

In contrast, the results of the critical knowledge dimension show a tendency for men to be more knowledgeable. Indeed, they self-reported greater knowledge (from 11 to 33 percentage points) in all the survey items in this section, with the greatest gap existing in how much the respondents claim to know about the basic features of the digital services they use.

4. Discussion and conclusions

The 16-to-18-year-olds in our sample belong to Generation Z, a population group that frequently uses a variety of technological devices every day, which may lead to the assumption that they possess digital skills. However, as Darvin (2018) points out, such skills cannot be preconceived in terms of age alone; and as Porat et al. (2018) note, previous studies have shown that time spent online is only relevant when it comes to technical skills.

According to our findings, although young people's self-perceived competence in many of the technical skills is high (above 70%), there are still certain actions they report finding difficult to carry out, which gives us an answer to [RQ1]. This is a key point considering Van-Deursen et al.'s (2021) observation that technical skills are essential, for it is impossible to browse the internet or use mobile devices without them. Regarding the young respondents' self-perceived lack of knowledge on how to bolster their privacy (only 27.8% know how to use tools to increase privacy and anonymity online), Van-Deursen et al. (2021) note that this is a widespread feeling amongst users in the Internet of Things era, as "the collection, analysis and use of collected data is often not transparent to users, making it more difficult to make decisions about whether or not to use a smart device" (5).

Overall, in terms of technical skills, there are no significant differences between young women and men, except for the greater competence self-reported by young men in repairing devices. However, as Weston et al. (2019) argue, the need to improve the technical digital skills of girls and young women should not be dismissed, as this increases the likelihood that they will go on to pursue higher education in the traditionally male-dominated fields of science, technology, engineering, and mathematics (STEM).

Differences in informational skills may be linked to the ways in which people perceive and learn to use the internet, which can differ between age groups and work environments. Previous studies have suggested that this link is stronger in older age groups than in younger groups (Van-Deursen et al., 2014). According to previous research, cultural background and level of education have a significant effect, which

is stronger in technical skills, while gender shows a more defined impact on theoretical knowledge (Gui & Argentin, 2011; Vod et al., 2022). When broken down by gender, our results show that informational skills seem to be fairly balanced overall, although some differences are observed as we go further into detail. In all, young Spaniards believe they have higher technical and informational digital skills than critical digital knowledge, which validates [H1].

Low proficiency in certain skills is consistent with De-Vicente-Domínguez et al. (2022), who found that being a digital native does not necessarily mean possessing digital skills, as many young people display a number of deficiencies, such as a lack of knowledge of Boolean expressions and search commands and operators, and unfamiliarity with meta-search engines and specific directories. The authors recommend that teachers "provide students with useful digital resources to access reliable content of the subject taught, thus expanding their skills to access digital culture" (De-Vicente-Domínguez et al., 2022: 152).

In terms of critical knowledge by gender, men are self-reportedly much more knowledgeable than women in all the survey items in this section, which validates [H2]. These results do not bear out an actual knowledge gap between young men and women in Spain, as the data are based on the respondents' self-perceptions. However, given the contrast between male "superiority" in these sections and the relative predominance of young women in the digital skills analysed, these findings can be interpreted from the perspective that technology, as a social construct, is historically and culturally associated with the male gender and plays a key role in power relations between genders, reproducing patriarchal logic (Wajcman, 2010). In this regard, our results highlight the need and relevance of developing digital literacy education programmes that integrate, at their core, the approaches of feminist theories that address the issue of technology, in order to close the gender digital divide (Gurung, 2018).

Although previous literature frames our analysis within the second level of the digital divide, this paper offers a first glimpse into the digital skills and knowledge of Spain's younger population (16-18 years), as well as a stepping stone towards the third level of the digital divide, which involves harnessing the potential of digital capital to provide resources and training. By breaking this third level down by gender, we raise the possibility of linking digital capital and political capital (Ragnedda, 2016) through a critical digital pedagogy that combines digital skills and knowledge aligned with techno-feminism (Wajcman, 2004) or cyber-feminism, especially applied to educational settings (Mérida, 2019), both for the use of digital media as an extension of feminist activism (Sánchez-Duarte & Fernández-Romero, 2017) and for politicising digital technologies from a feminist perspective (Binder & García-Gago, 2020).

"Feminist internet research" (Perera, 2022), when approached from an intersectional perspective, provides tools to tackle the previously mentioned problems facing the public sphere. With respect to dataveillance, it proposes a structural approach that goes beyond personal privacy, shedding light on how surveillance intensifies among oppressed groups (Kovacs, 2017). Meanwhile, it questions the knowledge produced by datafication, studies the effects of its unequal distribution of power, and advocates the creation of pluralist epistemologies and policies (D'Ignazio & Klein, 2020). Regarding artificial intelligence, big data and the algorithms that govern the current media ecosystem, this approach calls out their biases and consequent reinforcement of inequalities, while proposing AI governance models that correct these biases and promote algorithmic justice (Peña & Varon, 2019).

The results of this research have implications in the design of future educational programmes for young people at the intersection of digital technologies and gender equality, viewing the binary gender approach used here as a limitation to be overcome by addressing the issues raised by queer theory in the field of media literacy in the digital environment (Van-Leent & Mills, 2017). While gender equality has been addressed in the context of formal education (Prendes-Espinosa et al., 2020), as a future line of research we suggest exploring the enormous potential of non-formal education and social education to apply this perspective, given the dynamic and fluid nature of young people's digital practices. Here the focus should not be so much on technical skills, but on the need to develop informational skills and critical knowledge, in line with the findings of studies such as Porat et al. (2018), Martinovic et al. (2019) and Jackman et al. (2021). Education in this respect must be continuous and cross-cutting so that it can adapt to the constant changes in the digital environment and make the public more competent for future working environments, as well as more socially, culturally, and politically egalitarian and participatory.

Notes

[1]In our presentation of the survey results, we have reduced the Likert scale from five to three variables, merging the options "I know how to do this and could teach others" and "I know how to do this by myself" as "I know how to do this" on one side and the options "I know what this is but I don't know how to do it" and "I don't know what this is or what it means" as "I don't know how to do this" on the other. Figures under the "Total" heading do not add up to 100% because the tables do not include instances where the respondent said that they did not know how to answer or did not answer at all.

[2]Table 4 is attached as supplementary material.

Authors' Contribution

Idea, E.E., M.M., D.A.; Literature review (state of the art), P.F.; Methodology, E.E., M.M., L.M.; Data analysis, L.M.; Results, E.E., M.M., L.M.; Discussion and conclusions, E.E., M.M., P.F.; Writing (original draft), E.E., M.M., P.F.; Final revisions, E.E., M.M., P.F., D.A.; Project design and sponsorship, D.A.

Funding Agency

This article is a result of the research project "Educación social digital: juventud, ciudadanía activa e inclusión" [Digital social education: youth, active citizenship and inclusion] (reference: PGC2018-095123-B-I00), approved in the 2018 call for "Proyectos de I+D de Generación de Conocimiento" [Knowledge Generation R&D&I Projects] under the Programa Estatal de Generación de Conocimiento y Fortalecimiento Científico y Tecnológico del Sistema de I+D+i [National Programme for Knowledge Generation and for the Scientific and Technological Strengthening of the R&D&I System], funded by Spain's Ministry of Economy and Competitiveness.

References

Aranda, D., Sánchez-Navarro, J., & Mohammadi, L. (2020). Perception and self-assessment of digital skills and gaming among youth: A dataset from Spain. *Data in Brief*, *28*. https://doi.org/10.1016/j.dib.2019.104957

Bali, M. (2019). Reimagining digital literacies from a feminist perspective in a postcolonial context. *Media and Communication*, *7*(2), 69-81. https://doi.org/10.17645/mac.v7i2.1935

Binder, I., & García-Gago, S. (2020). *Politizar la tecnología. Radios comunitarias y derecho a la comunicación en los territorios digitales*. Ediciones Jinete Insomne. https://bit.ly/3wCVGNI

Choi, M. (2016). A concept analysis of digital citizenship for democratic citizenship education in the digital age. *Theory and Research in Social Education*, *44*, 565-607. https://doi.org/10.1080/00933104.2016.1210549

Cockburn, C. (1983). *Brothers: Male dominance and technological change*. Pluto Press. https://bit.ly/3NtLU67

Cohen, J.N. (2018). Exploring echo-systems: how algorithms shape immersive media environments. *Journal of Media Literacy Education*, *10*(2), 139-151. https://doi.org/10.23860/JMLE-2018-10-2-8

Darvin, R. (2018). Social class and the unequal digital literacies of youth. *Language and Literacy*, *20*(3), 26-45. https://doi.org/10.20360/langandlit29407

De-Vicente-Domínguez, A.M., Carballeda-Camacho, M.R., & Cestino-González, E. (2022). The analysis of the informational competencies of students entering university: a case study [Análisis de las competencias mediáticas del alumnado que ingresa en la universidad: Un estudio de caso en estudiantes de comunicación]. *Vivat Academia*, *155*, 151-171. https://doi.org/10.15178/va.2022.155.e1375

Díez-Gutiérrez, E.J. (2022). Concientización. In *Palabras y pedagogía desde Paulo Freire*. La Muralla. https://bit.ly/3vCodCf

D'ignazio, C., & Klein, L.F. (2020). *Data feminism*. The MIT Press. https://doi.org/10.7551/mitpress/11805.001.0001

Dubois, E., & Blank, G. (2018). The echo chamber is overstated: The moderating effect of political interest and diverse media. *Information, Communication & Society*, *21*, 729-745. https://doi.org/10-1080/1369118X.2018.1428656

Fernández-De-Castro, P., Aranda, D., Moyano, S., & Sampedro, V. (2021). *Digital youth work: A systematic review with a proposal*. Social Work Education. https://doi.org/10.1080/02615479.2021.1971187

Freire, P. (1970). *Pedagogy of the oppressed*. Continuum. https://bit.ly/3zZx3g0

Gleason, B., & Von-Gillern, S. (2018). Digital citizenship with social media: Participatory practices of teaching and learning in secondary education. *Educational Technology & Society*, *21*(1), 200-212. https://bit.ly/3S8D94B

Gui, M., & Argentin, G. (2011). Digital skills of internet natives: Different forms of digital literacy in a random sample of northern Italian high school students. *New Media & Society*, *13*(6), 963-980. https://doi.org/10.1177/1461444810389751

Gurung, L. (2018). The digital divide: An inquiry from feminist perspectives. *Dhaulagiri Journal of Sociology and Anthropology*, *12*, 50-57. https://doi.org/10.3126/dsaj.v12i0.22179

Gutiérrez-Martín, A., & Tyner, K. (2012). Media education, media literacy and digital competence. [Educación para los medios, alfabetización mediática y competencia digital]. *Comunicar*, *38*, 31-39. https://doi.org/10.3916/C38-2012-02-03

Hargittai, E. (2002). Second-level digital divide: Differences in people's online skills. *First Monday*, *7*(4), 1-19. https://doi.org/10.5210/fm.v7i4.942

Helsper, E.J. (2021). *The digital disconnect: The social causes and consequences of digital inequalities*. SAGE Publications Ltd. https://doi.org/10.4135/9781526492982

Hooks, B. (1994). *Teaching to transgress: Education as the practice of freedom*. Routledge. https://bit.ly/3Ntc4pG

Jackman, J.A., Gentile, D.A., Cho, N.J., & Park, Y. (2021). Addressing the digital skills gap for future education. *Nature Human Behaviour*, *5*(5), 542-545. https://doi.org/10.1038/s41562-021-01074-z

Kovacs, A. (2017). *Reading surveillance through a gendered lens: Some theory*. Internet Democracy Project. https://bit.ly/3BviMZs

Loh, J. (2019). What is feminist philosophy of technology? A critical overview and a plea for a feminist technoscientific utopia. In J. Loh, & M. Coeckelbergh (Eds.), *Feminist philosophy of technology* (pp. 1-24). J.B. Metzler. https://doi.org/10.1007/978-3-476-04967-4_1

Martinovic, D., Freiman, V., Lekule, C.S., & Yang, Y. (2019). The roles of digital literacy in social life of youth. In A. Ray (Ed.), *Advanced methodologies and technologies in library science, information management, and scholarly inquiry* (pp. 103-117). IGI Global. https://doi.org/10.4018/978-1-5225-7659-4.ch009

Mérida, J.A. (2019). Ciberfeminismo: Una perspectiva desde las aulas. *Íber*, *94*, 51-56. http://bit.ly/2mLZESV

Mihailidis, P., Shrestova, S., & Fromm, M. (2021). The values of transformative media pedagogies. In P. Mihailidis, S. Shrestova, & M. Fromm (Eds.), *Transformative media pedagogies*. Routledge. https://doi.org/10.4324/9781003031246-2

Pangrazio, L. (2016). Reconceptualising critical digital literacy. *Discourse: Studies in the cultural politics of education*, *37*, 163-174. https://doi.org/10.1080/01596306.2014.942836

Panke, S., & Stephens, J. (2018). Beyond the echo chamber: Pedagogical tools for civic engage¬ment discourse and reflection. *Educational Technology & Society*, *21*(1), 248-263. https://bit.ly/3vpDOoy

Peña, P., & Varon, J. (2019). Decolonising AI: A transfeminist approach to data and social justice. In A. Finlay (Ed.), *Global Information Society Watch 2019: Artificial intelligence: Human rights, social justice and development*. Association for Progressive Communications (APC). https://bit.ly/3oFDWfU

Perera, S. (2022). *White paper on feminist internet research*. Association for Progressive Communications. https://bit.ly/3MGFWyH

Porat, E., Blau, I., & Barak, A. (2018). Measuring digital literacies: Junior high-school students' perceived competencies versus actual performance. *Computers & Education*, *126*, 23-36. https://doi.org/10.1016/j.compedu.2018.06.030

Pötzsch, H. (2019). Critical digital literacy: Technology in education beyond issues of user competence and labour-market qualifications. *TripleC*, *17*(2), 221-240. https://doi.org/10.31269/triplec.v17i2.1093

Prendes-Espinosa, M., García-Tudela, P., & Solano-Fernández, I. (2020). Gender equality and ICT in the context of formal education: A systematic review. [Igualdad de género y TIC en contextos educativos formales: Una revisión sistemática]. *Comunicar*, *63*, 9-19. https://doi.org/10.3916/C63-2020-01

Ragnedda, M. (2016). *The third digital divide: A webberian approach to digital inequalities*. Routledge. https://doi.org/10.4324/9781315606002

Richardson, J.W., Martin, F., & Sauers, N. (2021). Systematic review of 15 years of research on digital citizenship: 2004-2019. *Learning, Media and Technology*, *46*, 498-514. https://doi.org/10.1080/17439884.2021.1941098

Sánchez-Duarte, J.M., & Fernández-Romero, D. (2017). Subactivismo feminista y repertorios de acción colectiva digitales: Prácticas ciberfeministas en Twitter. *Profesional de la Información*, *26*, 894-902. https://doi.org/10.3145/epi.2017.sep.11

Valtonen, T., Tedre, M., Mäkitalo, K., & Vartianen, H. (2019). Media literacy education in the age of machine learning. *Journal of Media Literacy Education*, *11*(2), 20-36. https://doi.org/10.23860/JMLE-2019-11-2-2

Van-Deursen, A.J.A.M., & Helsper, E.J. (2015). The third-level digital divide: Who benefits most from being online. In L. Robinson, S. R. Cotten, J. Schulz, T. M. Hale, & A. Williams (Eds.), *Communication and information technologies annual: Vol. 10*, volume 10 (pp. 29-52). Emerald Group Publishing Limited. https://doi.org/10.1108/S2050-206020150000010002

Van-Deursen, A.J.A.M., Helsper, E.J., & Eynon, R. (2014). *Measuring digital skills. From digital skills to tangible outcomes project report*. The London School of Economics and Political Science, University of Twente & Oxford Internet Institute University of Oxford. https://bit.ly/3NxnKYi

Van-Deursen, A.J.A.M., Helsper, E.J., & Eynon, R. (2016). Development and validation of the Internet Skills Scale (ISS). *Information, Communication & Society*, *19*, 804-823. https://doi.org/10.1080/1369118X.2015.1078834

Van-Deursen, A.J.A.M., Van-Der-Zeeuw, A., De-Boer, P., Jansen, G.P., & Van-Rompay (2021). Development and validation of the Internet of Things SkillsScale (IoTSS). *Information, Communication & Society*. https://doi.org/10.1080/1369118X.2021.1900320

Van-Dijck, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, *12*(2), 197-208. https://doi.org/10.24908/ss.v12i2.4776

Van-Leent, L., & Mills, K. (2017). A queer critical media literacies framework in a digital age. *Journal of Adolescent & Adult Literacy*, *61*(4), 401-411. https://doi.org/10.1002/jaal.711

Vod , A.I., Cautisanu, C., Gr dinaru, C., T n sescu, C., & De Moraes, G.H.S.M. (2022). Exploring Digital Literacy Skills in Social Sciences and Humanities Students. *Sustainability*, *14*. https://doi.org/10.3390/su14052483

Wajcman, J. (2004). *Technofeminism*. Polity Press. https://bit.ly/38Hhdvy

Wajcman, J. (2010). Feminist theories of technology. *Cambridge Journal of Economics*, *34*(1), 143-152. https://doi.org/10.1093/cje/ben057

Weston, T.J., Dubow, W.M., & Kaminsky, A. (2019). Predicting women's persistence in computer science-and technology-related majors from high school to college. *ACM Transactions on Computing Education (TOCE)*, *20*(1), 1-16. https://doi.org/10.1145/3343195

Yue, A., Nekmat, E., & Beta, A.R. (2019). Digital literacy through digital citizenship: Online civic participation and public opinion evaluation of youth minorities in Southeast Asia. *Media and Communication*, *7*(2), 100-114. https://doi.org/10.17645/mac.v7i2.1899

Comunicar, n. 74, v. XXXI, 2023 | Media Education Research Journal | ISSN: 1134-3478; e-ISSN: 1988-3478
www.comunicarjournal.com

Social media influence on young people and children: Analysis on Instagram, Twitter and YouTube

Redes sociales y su influencia en los jóvenes y niños: Análisis en Instagram, Twitter y YouTube

Dr. Raquel Lozano-Blasco. Associate Professor, Department of Psychology and Sociology, University of Zaragoza (Spain) (rlozano@unizar.es) (https://orcid.org/0000-0002-0100-1449)
Marta Mira-Aladrén. Research Trainee, Department of Psychology and Sociology, University of Zaragoza (Spain) (mmira@unizar.es) (https://orcid.org/0000-0002-6088-0324)
Mercedes Gil-Lamata. Research Trainee, Department of Business Management and Organisation, University of Zaragoza (Spain) (mercedesgil@unizar.es) (https://orcid.org/0000-0002-1816-5394)

ABSTRACT

Social networking sites are a new ecosystem of social relations in which adolescents follow public figures or influencers: instagrammers, tweeters and youtubers. Their behaviour in the posts they publish become a trend and a model for the new generations. In order to explore these behaviours and their consequences, it is useful to study the behaviour of the 10 instagramers, 10 tweeters and 10 youtubers with the largest number of followers in the world. A mixed method was employed, combining: social media analysis (SNA) methodology executed by monitoring Twitter, Instagram and YouTube accounts and their publications (300 posts with the highest number of likes). The FanapageKarma tool was used to capture data by applying data mining techniques. Subsequently, sentiment analysis was performed using Meaning Cloud software, determining sentiment polarity analysis quantitatively. Finally, a semantic analysis of the content was performed using Nvivo. The results of multi-regression and sentiment's analysis show clear differences between social networking sites. Twitter is a space for critical analysis of information and social movements, especially climate change. In this space adolescents defend their values and ideology. Instagram is a showcase for fashion and beauty, where brands support an idealised and desirable lifestyle. YouTube is a space for entertainment and comedy. It concludes that despite their differences there is one univocal feature, the effort of influencers to capture audiences and establish parasocial relationships.

RESUMEN

Las redes sociales son un nuevo ecosistema de relaciones sociales en el que los adolescentes siguen a personajes públicos o «influencers»: «instagramers», «twitteros» y «youtubers». Su comportamiento en los posts que publican se convierte en una tendencia y un modelo para las nuevas generaciones. Para profundizar en estos comportamientos y sus consecuencias, resulta de utilidad estudiar el comportamiento de los 10 «instagramers», 10 «twitteros» y 10 «youtubers» con mayor número de seguidores en el mundo mediante sus publicaciones (300 post con mayor cantidad de likes). Se empleó un método mixto, combinando: la metodología de análisis de medios sociales (SNA) ejecutada mediante la monitorización de cuentas de Twitter, Instagram y YouTube. Se empleó el instrumento de FanapageKarma para captar los datos aplicando técnicas de minería de datos. Posteriormente, se realizó un análisis de sentimiento mediante el software «Meaning Cloud», este determinó el análisis de la polaridad de los sentimientos de forma cuantitativa. Finalmente, se realizó un análisis semántico de los contenidos mediante Nvivo. Los resultados de la multirregresión y el análisis de sentimientos muestran claras diferencias entre las redes sociales. Twitter es un espacio de análisis crítico de la información y de los movimientos sociales, especialmente del cambio climático. En este espacio los adolescentes defienden sus valores e ideología. Instagram es un escaparate de moda y belleza, donde las marcas apoyan un estilo de vida idealizado y deseable. YouTube es un espacio para el entretenimiento y la comedia. Se concluye que a pesar de sus diferencias hay una característica unívoca, el esfuerzo de los «influencers» por captar audiencias y establecer relaciones parasociales.

KEYWORDS | PALABRAS CLAVE

Adolescence, youth, polarity, Twitter, YouTube, Instagram.
Adolescencia, juventud, polaridad, Twitter, YouTube, Instagram.

Received: 2022-05-02 | Reviewed: 2022-06-13 | Accepted: 2022-07-19 | OnlineFirst: 2022-10-30 | Published: 2023-01-01
DOI https://doi.org/10.3916/C74-2023-10 | Pages: 117-128

1. Literature review

The establishment of social networks has meant the structuring of virtual spaces where adolescents relate and share interests and experiences (Burnette et al., 2017), meeting different social needs (Anderson & Jiang, 2018; Smith & Anderson, 2018; Throuvala et al., 2019; Vannucci & McCauley-Ohannessian, 2019).

Instagram is a space where social comparison takes place through a canonized model of beauty, that is, it operates an idealised comparison (Burnette et al., 2017; Kim & Kim, 2019; Schmuck et al., 2019; Yau & Reich, 2019) that models self-concept (Kim et al., 2017; Verrastro et al., 2020). This situation is propitious for the emergence of the well-known Instagrammer, users whose function is to influence the virtual community by showing aspects of their lives and establishing themselves as marketing images of the companies that sponsor them (Boerman, 2020; Schouten et al., 2020; Weismueller et al., 2020). The linguistic simplicity and strength of the image means that it is preferred by influencers to capture an audience (Shane-Simpson et al., 2018).

On the other hand, YouTube is a social network based on audio-video content in which videos about people's personal lives and interests (Jerslev, 2016) or celebrities (Ferchaud et al., 2018; Hartmann, 2016) are regularly presented. In other words, YouTubers are established as video bloggers who regularly post videos about their lives, interests or skills, propose challenges and participate in conversations with other users, i.e. a parasocial relationship is generated between influencers and followers (Ashman et al., 2018; de-Bérail et al., 2019; Genç & Öksüz, 2019). Nevertheless, a study with Spanish teenagers has argued that the most appreciated values are humour and leisure, that is to say, these users value their comedy and not the image of a brand that they can represent (Aran-Ramspott et al., 2018). Twitter is considered the network par excellence of information and communication, being used to debate social, environmental and political issues, rendering it a thermometer of social concerns (Harb et al., 2020; Neu et al., 2019; Peres et al., 2020).

In short, teenage influencers are teenagers who narrate and publicly show their lives (fashion looks, romantic stories, make-up skills, socio-political concerns, etc.) (Ferchaud et al., 2018; Hartmann, 2016) through social networks, and can consequently be considered child and youth celebrities (Aran-Ramspott et al., 2018). In most cases, they do not present a transgressive vision but imitate sponsors in the publications they share (Blasco-García, 2020). Their ability to reach out to the teenager gives them a privileged position from a marketing standpoint (Bakir et al., 2020; Boerman, 2020; Schouten et al., 2020). In this way, they favour the transmission of social norms based on consumerism, which not only affects the target population but also normalizes cruel optimism (Ashman et al., 2018). On the other hand, a dichotomous parasocial relationship is established between influencers and followers, so that the latter empathise to such an extent that a physical encounter is not necessary (Ashman et al., 2018; de-Bérail et al., 2019; Genç & Öksüz, 2019). For instance, Bhatia (2018) shows how YouTubers 'discursively exploit the boundaries', oscillating between expert and colloquial knowledge and taking advantage of their discursive competence. In this way, they show themselves to be participatory and committed, making users feel part of the YouTube community as they address them, surely read their comments on videos and so forth, thereby generating an emotional commitment (Bhatia, 2018; Burgess & Green, 2009; Lange, 2014). Specifically, Bhatia's (2018) results show how YouTubers carefully prepare their videos, presenting some degree of strategy. Scannell (2000) argues that conversational quality is a determining factor in sociability. In short, publications on social networks are far from natural; on the contrary, they show a communicative strategy.

From this perspective, we should ask ourselves why adolescents use social networks. Existing research findings suggest that adolescents engage in social sharing and self-editing to adjust their views of themselves to their ideal self. In other words, they seek idealised self-representation (Burnette et al., 2017; Schmuck et al., 2019; Stockdale & Coyne, 2020). This pursuit of the ideal can reinforce low self-esteem and low self-concept and even instigate risky eating behaviours (Burnette et al., 2017; Schmuck et al., 2019), although authors such as Stockdale and Coyne (2020) highlight how, in addition to social comparison, boredom plays a key role. Authors such as Mäntymäki and Riemer (2014) argue that social network use is mediated by hedonism, i.e. a search for pleasure and satisfaction in an online life. Others point out that it is the result of a sense of belonging and self-disclosure, key processes in identity construction at this stage (Davis,

2012). While the theory of uses and rewards and the theory of self-determination expose the relevance of controlling relationships, content, presentation and impressions, this may be the same etiology of fear of missing out (FOMO) and nomophobia (Throuvala et al., 2019). On the other hand, it is necessary to recognise how authors like de-Bérail et al. (2019) and Hartmann (2016) are based on the theory of parasocial relations (which try to satisfy those social needs that are lacking). In this sense, people with anxiety, loneliness or social difficulties rely on parasocial relationships to compensate for their difficulties. Consequently, a bond is generated between the influencer and their followers.

In terms of coherence, the state of the art shows univocal traits in terms of polarity of feeling in social networks, with the idealised vision of oneself and any situation being practically hegemonic. Thus, positive or neutral texts with a high degree of subjectivity abound (Peres et al., 2020; Reyes-Menéndez et al., 2018; Vizcaíno-Verdú & Aguaded, 2020). However, there are certain dichotomies according to the subject matter, so that body image, self-expression, travellers, digital life and startups are associated with positivity, while those related to depression, loneliness and real-world relationships have negative polarity and self-identity and anxiety are neutral (Saura et al., 2019).

One of the outstanding features of social network publications is the use of the hashtag, a short fragment headed by # which manages to increase a publication's visibility and virality (Lipsman et al., 2012; McGoogan, 2017). Erz et al. (2020) show how the use of hashtags correlates with a person's idealized exposure. Similarly, the presentation of consumer elements at the beginning of YouTube videos increases their market share (van-Reijmersdal et al., 2020). As for the most followed content on YouTube, Castillo-Abdul et al. (2020) refer to important gender differences, in such a way that boys stand out as 'gamers' while girls do so in lifestyle issues.

2. Research questions and objectives

It is important to examine the behaviour of adolescents in partner networks by studying their relationships with influencers. The following research questions are thus posed: Which posts do adolescents value most from their influencers? Are there differences according to the social network? What type of content is best received by the adolescent population?

The general objective of the research is to understand the behaviour of the most relevant influencers for the youth population in three of the main social networks. To this end, three specific objectives will be developed: to identify the most influential accounts at present; to carry out a sentiment analysis (polarity, agreement, irony, subjectivity) of the main contents of their publications with the greatest impact (number of likes); and to study the differences in the publications of the aforementioned accounts depending on the social network. Accordingly, the following hypotheses are proposed:

- H1. Within the widespread use of social networks among children and young people, YouTube and Instagram influencers attract more followers than Twitter influencers.
- H2. The content that most attracts teenagers differs according to the social network and its function. Twitter has an information and social awareness function, Instagram a space to present a self-idealised image of oneself, and YouTube playful and entertaining.
- H3. The content with the greatest impact finds justification in the theory of social comparison and social relations.

3. Dataset(s) and methods

3.1. Sample

The corpus consisted of 100 posts of Instagram, 100 tweets, and 100 YouTube videos, extracted from the accounts of international child and youth influencers with the largest numbers of followers, comprising 10 Tweeters, 10 Instagrammers and 10 YouTubers (Table 1). In this sense, the 10 publications with the highest number of likes from each influencer were chosen. To select the sample, the following were used: the platforms Hype Auditor (Instagram and YouTube) and Statista (Twitter), which provide information on the infant-juvenile influencers with the largest numbers of followers after applying the inclusion and exclusion criteria. However, it was necessary to apply inclusion and exclusion criteria to find child and youth influencers. The inclusion criteria were: a) accounts aimed at children and young

people (the contents and topics worked on must be interesting for this age group and follow digital culture trends); b) content with minors (where the protagonists are children or teenagers); c) minors regarded as influencers (these channels must have a large number of followers). And, as exclusion criteria; d) family accounts of social networks (families that present their daily life, adventures and misadventures of their family nucleus with traditional content); and e) accounts aimed at educational content (accounts aimed at language learning, music, autonomy, such as Cantajuegosvevo). In this sense, today's child and youth influencers belong either to generation Z or centennials (born between 1994 and 2010) or to generation Alpha (born between 2010 and 2025).

Table 1. Channels in each social network

Channel	Fans	Number of likes	Gender	Year of birth	Generation	Language	Country
Instagram							
@alaia	832,895	557,498	Female	2015	Alpha	Spanish	Spain
@clementstwins	1,827,213	9,511,370	Female	2010	Z	English	USA
@brynnrumfallo	3,123,885	8,457,731	Female	2003	Z	English	USA
@coco_pinkprincess	641,995	721,706	Female	2010	Z	English/ Japanese	Japan
@daniellecohn	4,765,218	43,523,845	Female	2004	Z	English	USA
@Gibby :)	1,604,447	1,347,167	Female	2007	Z	Spanish	Mexico
@fashion_laerta	1,145,373	790,700	Female	2011	Alpha	English	UK
@Melody	11,100,000	62,800,834	Female	2007	Z	Portuguese	Brazil
@Анастасия Князева	1340218	3,849,744	Female	2011	Alpha	Russian	Russia
@milliebellediamond	1,008,344	4,088,649	Female	2014	Alpha	English	Australia
Twitter							
@GretaThunberg	4,162,102	6,407,378	Female	2003	Z	English	Sweden
@heybalaa	63,990	2,959,229	Female	2004	Z	English	Mexico
@LittleMissFlint	128,529	1,515,814	Female	2007	Z	Spanish	USA
@marsaimartin	127,756	680,246	Female	2004	Z	English	USA
@Milliestopshate	590,195	884,225	Female	2004	Z	English	UK
@Estilosophie	62,823	526,072	Female	2005	Z	Spanish	Colombia
@ThiagolUTU	120,150	24,138	Male	2008	Z	Spanish	Argentina
@itslittlevale	412,479	6,130	Female	2002	Z	Spanish/English	Colombia
@ximeponchof	189,849	33,171	Female	2002	Z	Spanish	Mexico
@QueAmara	848,199	19,006	Female	2003	Z	Spanish	Colombia
YouTube							
@Bratayley	14,400,000	3,927,915	Female	2004	Z	Spanish	Spain
@Diana and Roma ESP	13,600,000	4,820,898	Female	2014	Alpha	Spanish/English	USA
@Evan-Tube-HD	11,600,000	2,067,844	Male	2005	Z	English	USA
@Gibby :)	11,300,000	716,823	Female	2007	Z	Spanish	Mexico
@Its JoJo Siwa	11,200,000	1,048,033	Female	2003	Z	English	USA
@MELODY OFICIAL	6,280,000	2,605,115	Female	2007	Z	Portuguese	Brazil
@Mimi Land	5,640,000	1,136,267	Female	2008	Z	English/Spanish	USA
@SIS vs BRO	4,390,000	2,734,924	Female and Male	2007 and 2008	Z	English	Canada
@Tiana Wilson	4,050,000	968,501	Female	2007	Z	English	UK
@Tiana	4,000,000	558,179	Female	2007	Z	English	UK

3.2. Data acquisition and procedure

The first part of the study was carried out using the social media analysis (SNA) methodology based on data mining. The data acquisition was done using Fanpage Karma software from 8 January 2020 to 8 September 2020. We collected 300 posts with the highest numbers using the tool Fanpage Karma. We followed the indications of previous research, such as Latorre-Martínez et al. (2018) and Lozano-Blasco et al. (2021). In this sense, the 10 posts with the highest number of likes from each influencer were selected. In this way, the 100 posts from each social network are proportional. This generated a considerable volume of data, materialized as key performance indicators (KPIs): number of likes, number of retweets, commitment and number of fans (Keegan & Rowley, 2017).

The second part of the study was carried out via opinion mining methodology based on the recognition of linguistic patterns through algorithms (Nguyen & Le-Nguyen, 2018). In other words, a sentiment analysis (Hu & Liu, 2004) of the 300 publications was conducted. Sentiment analysis analyzes the emotional character of the messages emitted from natural language, providing a holistic vision of the new ecosystem generated in social networks (Du cu & Günneç, 2020; Oramas-Bustillos et al., 2019; Yu et al., 2013). This was executed using the MeaningCloud tool and the Emotion Recognition pack, allowing it to be examined in several languages, in response to the linguistic diversity encountered. The third part of the research corresponded to a qualitative methodology, in which a semantic analysis of the publications was

carried out. Words with more than four letters were selected in order to avoid semantic categories such as articles, pronouns and prepositions, instead favoring the appearance of nouns, adjectives and verbs, because these reflect, to a greater extent, the complexity of speech (Krippendorff, 1980). In addition, this selection allowed us to recognize the appearance of "hashtags". In the same way, we operated according to synonyms by grouping words according to their meaning. This section was carried out by means of the qualitative software NVivo.

3.3. Instruments

"Fanpage Karma": this application allowed web analysis and the monitoring of partner networks such as Instagram, Twitter and YouTube, allowing the study of consumer trends. This platform has been used to capture data from the main channels of child and youth influencers (Keegan & Rowley, 2017).

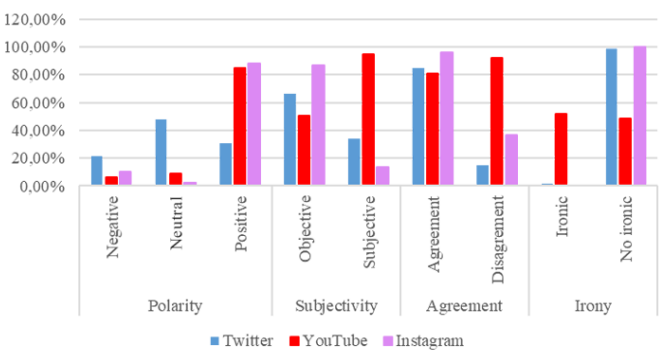
MeaningCloud Emotion Recognition Pack: this application was used in data mining to perform sentiment analysis through semantic recognition by algorithms. It is based on Robert Plutchik's theory of the 'emotion wheel'. In addition, it allows text input in different languages, making it possible to cater for linguistic diversity. The variables it studies are polarity, agreement, irony, subjectivity and confidence for polarity analysis. The reliability of this software is presented in previous research (Sharma & Hoque, 2017; Song et al., 2022; Xu et al., 2022). These variables were categorized and worked on according to the principles of Barton and Lazarsfeld (2015), Kale and Jayanth (2019), and León and Montero (2015).

NVivo: this is specialized qualitative research software. This one is specialized in making semantic analysis of content, generating word frequency and clouds of branched graphic words. It also supports different information, including text (Twitter), video and audio (YouTube) and text and image (Instagram).

4. Analysis and findings

The results returned by MeaningCloud expose diversity in the polarity of feelings among social networks (Figure 1).

Figure 1. Polarity of the three social networks



The network with the most critical language is Twitter, presenting a higher rate of negative than positive polarity, in contrast to Instagram and YouTube, where the opposite occurs. While Instagram's position has some similarity to Twitter, 85% of YouTube's content was categorized as positive, with residual negative content. As for the other variables in terms of sentiment analysis (Figure 1), the confidence index on polarity, although high, varies according to the network: Instagram registered 99.51%, Twitter 98.57% and YouTube 80.98%. As for the remainder of the variables, there are differences. Subjectivity indicates that although the most followed content on Instagram and Twitter is of an objective nature, that is, describing a fact or product, YouTube has a high rate of subjectivity, that is, the most followed content is personal opinion. YouTube consistently presents a higher level of affectivity (disagreement) than Twitter and Instagram, that is, its language is loaded with emotion, especially pleasant feelings. Similarly, it is unsurprising that YouTube also presents a higher rate of irony, ironic content on Twitter and Instagram

being minimal. Consequently, H1 is accepted, as YouTube and Instagram are more active than Twitter among the younger generation.

Pearson's correlations show strong and moderate relationships between the sentiment analysis variables of the three social networks (Table 2). However, simple forward regression tests on polarity reveal that not all variables have the same weight on the three social networks (Table 3).

Table 2. Correlations: Social networking sites

	1. Polarity	2. Agreement	3. Subjectivity	4. Irony
Instagram				
1. Polarity	1	.167**	.481**	.461**
2. Agreement		1	.168**	.627**
3. Subjectivity			1	.544**
4. Irony				1
Twitter				
1. Polarity	1	.297**	.571**	.134*
2. Agreement		1	.156**	.203**
3. Subjectivity			1	.215**
4. Irony				1
YouTube				
1. Polarity	1	.664**	.874**	.266**
2. Agreement		1	.609**	.305**
3. Subjectivity			1	.237**
4. Irony				1

Initially the polarity of Instagram is explained by 27.1% (R2=0.271, p<.000), the role of subjectivity and the level of confidence being statistically significant. That is, the number of terms referring to pleasant or unpleasant feelings is explained by the degree of subjectivity (writing of a fact or product, or personal vision or opinion) used in a message, together with the level of confidence in the success of the polarity.

Similarly, the polarity of Twitter is explained by 39% (R2=0.390, p<.000), by the level of subjectivity and the level of trust. Although YouTube shows different behaviour, its polarity is explained in 79.3% (R2=0,739, p<.000) by the agreement variable and subjectivity. That is, its polarity is determined by the number of emotionally charged words and the use of terms that denote a personal opinion or vision (Table 3).

Table 3. Polarity regressions (dependent variable) according to networks

Social networking sites	B	Standard error	Beta	t	Sig.	R2
Instagram						
(Constant)	11.629	10.584		1.099	0.273	0.271***
Agreement	-1.227	0.792	-0.118	-1.549	0.122	
Subjectivity	2.456	0.270	0.459	9.085	0.000	
Twitter						
(Constant)	-8.245	5.690		-1.449	0.148	0.390***
Agreement	0.296	0.393	0.052	0.755	0.451	
Subjectivity	2.126	0.201	0.501	10.577	0.000	
YouTube						
(Constant)	-26.283	1.342		-19.578	0.000	0.793***
Agreement	0.597	0.121	0.187	4.957	0.000	
Subjectivity	2.931	0.136	0.738	21.621	0.000	

Note. Level of significance p<0.001.

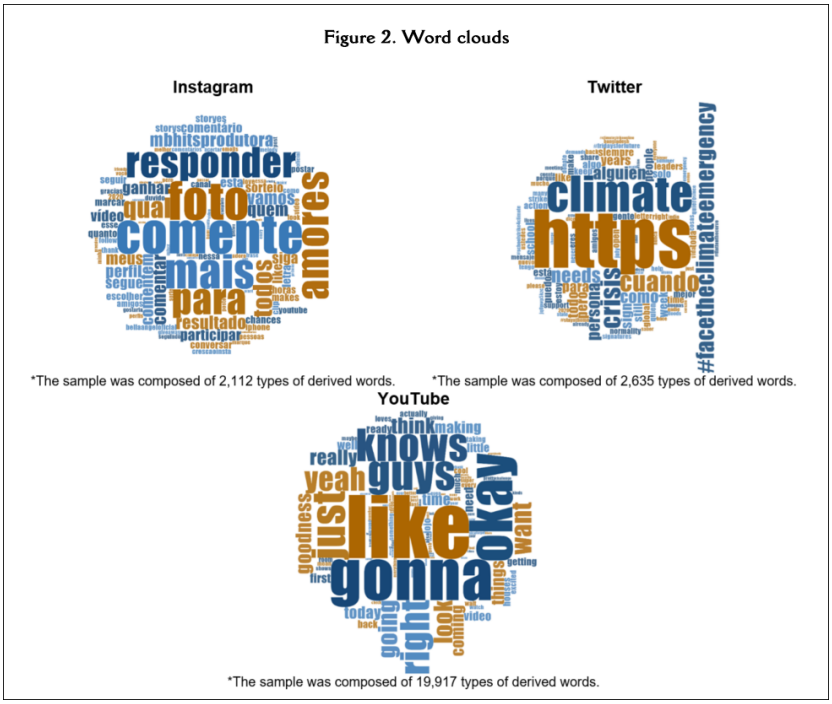
As for the semantic analysis of the language most used on social networks, it can be said that the total sample was composed of 24,664 words in English, Spanish, Portuguese and Russian. The search for so-called stop-words was applied automatically. These words are recognised as meaningless (regardless of the number of letters they have) and must be eliminated before the semantic study.

The distribution of the most used words is a function of the level of density, generating word clouds (Figure 2). In this sense, if we analyse in depth the most frequent terms, we find differences in the themes of each social network. On the one hand, Instagram exposes a universe of terms aimed at stimulating interaction with users' followers: "answer", "comment", "love", "follow", "like", "photo". In this way, users seek to promote influence among their followers, demanding attention and greater interaction while presenting themselves as subjects' worthy of following.

As far as Twitter is concerned, the most used terms are sustainability and adjectives that express social movements or actions, such as "climate", "crisis", "needs" and "person". Two hashtags even appear:

"#facetheclimateemergency" and "#flattenthecurve". The use of 'https' is also surprising, as it shows how Twitter is not only a text space but also a video one. Finally, YouTube exhibits behaviour similar to that of Instagram, in that the most used words refer to a conversation between a YouTuber and their followers, in which the former repeatedly launches terms to maintain attention: "cool", "look", "just", "really". In other words, influencers respond to a monologue-like narrative in which the sender refers to their audience: the content most supported by the adolescent population varies according to social network. In short, H2 is fulfilled, as the three networks studied manifest differences in their most-shared content, Instagram being a space of self-representation, Twitter of awareness of social movements and ecology and YouTube of purely playful entertainment.

Figure 2. Word clouds



On the other hand, the stream graph (Figure 3) shows the behaviour of the five most-used words in the 20 publications of greatest impact. Instagram shows how the terms "amores" (Spanish language, the translation is loves) and "vou" (Portuguese language, the translation is you), which are both used to address followers, remain constant and are interrelated with the words "comment", "reply" and "photo". In other words, the publications that teenagers like the most have as their backbone those that ask for help to increase the Instagrammers' own popularity.

On Twitter, there is a thematic line related to social and political actions ("leaders") to stop climate change – "climate", "crisis", "#facetheclimateemergency" – mostly accompanied by videos and other "https" links. As far as YouTube is concerned, there is a strong degree of homogeneity, with the terms "like", "gonna", "okay", "just" and "I'm' succeeding each other in a regular way". In other words, a narrative is presented in which the YouTuber interacts and seeks the complicity of their followers. In short, as in the case of Instagram, the publications with the greatest number of likes are those in which the influencers address their followers directly, seeking their direct interaction. Thus, H3 is fulfilled: the contents with the most impact are those where a parasocial relationship is established to influence followers.

In terms of hashtags, there are important differences between Twitter and Instagram. On Twitter there are two main topics: the school climate strike ("#FaceTheClimateEmergency" (30%), "#climatestrikeonline" (30%), "#fridaysforfuture" (7.50%), and "#schoolstrike4climate" (6.25%)) and COVID-19 and the social actions resulting from it ("#flattenthecurve" (6.25%), "#StayAtHome" (6.25%)). These two themes are strongly connected with the current social reality.

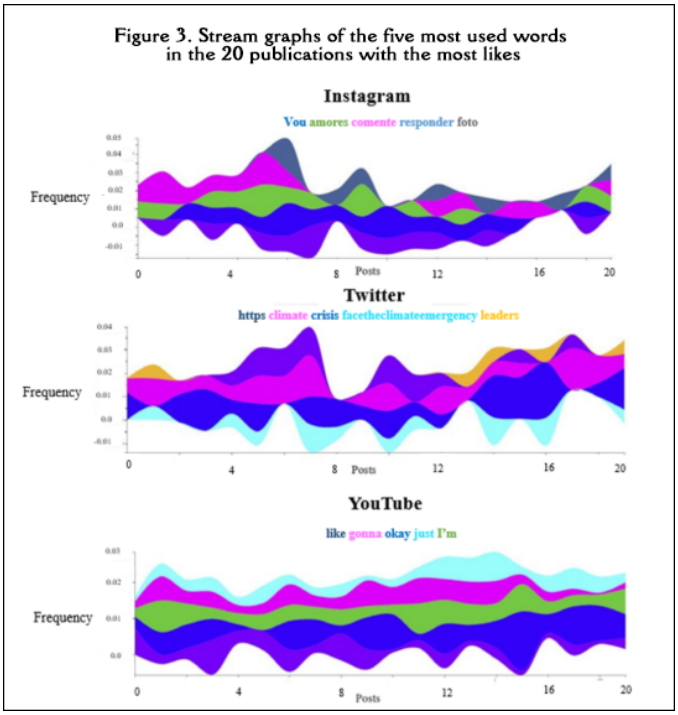


Figure 3. Stream graphs of the five most used words in the 20 publications with the most likes

Even though Instagram presents greater variety and dispersion, it is striking how a significant percentage refers to brands: "#SHEIN" (12.50%), "#nyxcosmetics_en" (9.37%), "#Zalan-doStyle" (6.25%), "#LGG8X" (3.12%) and "#APPLE" (3.12%). In this case, main Instagram topics are linked with marketing of fashion and technology brands.

5. Discussion and conclusions

Social networks operate within adolescents' everyday lives. The establishment of parasocial relationships between followers and influencers (Instagrammers, YouTubers and Tweeters) is of considerable importance for teenagers today. The differences between the social networks Instagram, Twitter and YouTube operate according to the different functions they perform, the results being consistent with previous publications (Throuvala et al., 2019; Vannucci & McCauley-Ohannessian., 2019).

Although all social networks are regularly used by the adolescent community (Anderson & Jiang, 2018; Smith & Anderson, 2018), the level of active participation seems to be mediated by the social network in question. In this regard, the results of fans and number of likes agree with previous studies indicating that YouTube is the most used platform among this population (Anderson & Jiang, 2018; Ofcom, 2017).

With respect to the sentiment analysis, similar results to previous investigations have been obtained. The contents of the study have considerable positivity or neutrality while also presenting significant subjectivity (Peres et al., 2020; Reyes-Menendez et al., 2018; Vizcaíno-Verdú & Aguaded, 2020). However, it should be noted that there are differences among the social networks, with Twitter being the most critical and objective and YouTube being the most positive and subjective. The results of the correlations and regressions specify how the polarity presents disparate actions. In the cases of Instagram and Twitter, the polarity is explained by "subjectivity" in moderate percentages. By contrast, the polarity of YouTube is the result of "agreement" and "subjectivity" in a high percentage. In summary, the sentiment analysis shows that, in the case of YouTube, messages are charged with "positive" emotions and personal views, as opposed to Twitter and Instagram.

On the other hand, we agree with Saura et al. (2019) that posts with a "real-world relationship" theme, especially those related to environmental policy violations, have a strong negative charge. In the same way, the categories body image, self-expression, travellers, digital life and startups have a positive or very positive polarity, being common in videos about the YouTubers' lifestyles or in Instagram publications.

Similarly, the semantic analysis of the word densities (word clouds), stream graphs and hashtags most used in publications with greater impact, has revealed significant differences among the social networks. Firstly, we agree with Blasco-García (2020) that posts do not present a transgressive image, but follow the fashion of the moment. It is necessary to make an important reflection on the use of the hashtag. The use of this technique increases visibility and virality, as can be seen in the results of this research, in line with previous studies (Lipsman et al., 2012; McGoogan, 2017).

However, there are profound differences between Twitter and Instagram in terms of the use of hashtags: whereas on Twitter its use responds to a social movement, on Instagram it pertains to fashion, cosmetics and technology brands, as also noted in previous research (Bakir et al., 2020; Boerman, 2020; Schouten et al., 2020). Indeed, Instagram is much more than a mirror to proclaim an ideal self, but rather a platform to engage an audience, based on communication that encourages interaction: "respond, picture, comment, love". In other words, the results are consistent with Shane-Simpson et al. (2018), who demonstrate how this social network is used to capture an audience, while its outreach to adolescents renders it a marketing tool (Bakir et al., 2020; Boerman, 2020; Schouten et al., 2020).

Besides that, if interest is focused exclusively on word density (word clouds) and stream graphs, similar communicative features can be identified, although with different themes. At first glance, one can note how Instagram refers to a world of self-representation, where social comparison is generated through selfies, attending to a canon of beauty (Burnette et al., 2017; Schmuck et al., 2019; Verrastro et al., 2020; Yau & Reich, 2019). Followers show their emotional attachment to publications in which Instagrammers ask them to collaborate or increase their influence. On another note, Twitter deals with social changes and actions, which turn out to be the thermometer of social concerns in the case of adolescents, for instance climate change (Neu et al., 2019; Peres et al., 2020). Teenagers present the social value generated by this age group as well as their ability to reason and defend their ideas and values. YouTube collects action verbs needed for humorous storytelling by both gamers and lifestyle bloggers, consistent with Aran-Ramspott et al. (2018) and Castillo-Abdul et al. (2020). The most relevant publications are those in which the YouTuber asks for attention and collaboration. The use of this type of expression coincides with the results of Bhatia (2018) and Scannell (2000) in such a way as to show a communicative strategy to achieve commitment and active participation. The use of terms that facilitate the establishment of communication among Instagrammers ("respond", "comenten" (translation comment), "loves", "photo", "vou" (translation you) and YouTubers ("I'm", "gonna", "okay", "just", "like") responds to the need to establish a parasocial relationship.

In this way, an affective bond and emotional commitment is generated between influencer and follower (Bhatia, 2018), which is rewarded by followers giving "likes". In this sense, these results coincide with previous studies that either explain this relationship or the extent to which these influencers normalise certain behaviours (Ashman et al., 2018; de-Bérail et al., 2019; Hartmann, 2016). In other words, the publications with the greatest scope are those in which a celebrity seeks their followers' participation, which, in the eyes of authors such as de-Bérail et al. (2019) and Hartmann (2016) could be indicative of high rates of loneliness and social difficulties, leading them to seek refuge in parasocial relationships.

This study is not without its limitations. First, the context of the pandemic must be taken into account, as this element may have increased connectivity for adolescents. On the other hand, it was not possible to find studies with which to make a true comparison of the results due to the speed with which new communication mechanisms and platforms are incorporated. On the other hand, the tools used are constantly changing and developing. Therefore, it is necessary to replicate this research in the coming years with more up-to-date and cutting-edge software. Similarly, it is necessary to generate longitudinal studies involving the study of a larger number of publications. Finally, given the pseudonymisation and the margin for lies inherent to social networks, bigdata research in this field has a limitation in the study of attitudes. This is because we will never be able to access what these people really think and feel, but only their behaviour in networks. For this reason, it seems interesting to carry out a study of the behaviour of influencers and their followers on all their profiles on all the social networks they have, including emerging channels such as TikTok or Twitch. In this way, it would be possible to analyse whether it is a behaviour adapted to the characteristics of the network used, such as market demand, or a behaviour maintained

in networks and over time, more similar to a study of attitudes. It would be interesting to carry out this study in a few years' time and observe changes in behaviour between today's adolescents and those in the future: will face-to-face social relationships have disappeared? Did social rules and social behaviour change in adolescents? Is the concept of friendship changing in adolescence? Do adolescents feel lonely and therefore establish parasocial relationships through social networks? Are adolescents' capacities for social relationships diminishing? What role do social networks such as Instagram play in marketing and attracting teenagers as customers? And as advertisers or marketers? Will new behaviours emerge linked to new social networks, with new features, such as TikTok or Twitch? As for the direct implications of this research, it is necessary to highlight how it has allowed us to understand what aspects are most valued by the digital community of children and youth influencers. Likewise, the behavior of influencers in this age group is different depending on the social networking sites used.

Authors' Contribution

Idea, R.L.B, M.M.A, M.G.L; Literature review (state of the art), R.L.B, M.M.A, M.G.L; Methodology, R.L.B, M.M.A, M.G.L; Data analysis, R.L.B, M.M.A, M.G.L; Results, R.L.B, M.M.A, M.G.L; Discussion and conclusions, R.L.B, M.M.A, M.G.L; Writing (original draft), R.L.B, M.M.A, M.G.L; Final revisions, R.L.B, M.M.A, M.G.L; Project design and funding agency, R.L.B, M.M.A, M.G.L.

Funding Agency

Regional Government of Aragón with a grant for the recruitment of pre-doctoral research staff in training for the period 2020-2024, and Spanish Ministry of Universities (SIA: 998758 2019).

References

Anderson, M., & Jiang, J. (2018). Teens, social media & technology. *Pew Research Center*. https://pewrsr.ch/3aRyOSL

Aran-Ramspott, S., Fedele, M., & Tarragó, A. (2018). YouTubers' social functions and their influence on pre-adolescence. [Funciones sociales de los Youtubers y su influencia en la preadolescencia]. *Comunicar*, *57*, 71-80. https://doi.org/10.3916/C57-2018-07

Ashman, R., Patterson, A., & Brown, S. (2018). 'Don't forget to like, share and subscribe': Digital autopreneurs in a neoliberal world. *Journal of Business Research*, *92*, 474-483. https://doi.org/10.1016/j.jbusres.2018.07.055

Bakir, A., Gentina, E., & De-Araújo-Gil, L. (2020). What shapes adolescents' attitudes toward luxury brands? The role of self-worth, self-construal, gender and national culture. *Journal of Retailing and Consumer Services*, *57*. https://doi.org/10.1016/j.jretconser.2020.102208

Barton, A.H., & Lazarsfeld, P.F. (1955). *Some functions of qualitative analysis in social research*. Bobbs Merrill.

Bhatia, A. (2018). Interdiscursive performance in digital professions: The case of YouTube tutorials. *Journal of Pragmatics*, *124*, 106-120. https://doi.org/10.1016/j.pragma.2017.11.001

Blasco-García, J. (2020). *Nuevas formas de ausencia: Las redes sociales*. [Doctoral Dissertation, Universitat Politécnica de Valencia]. https://bit.ly/3JplAJN

Boerman, S.C. (2020). The effects of the standardized Instagram disclosure for micro- and meso-influencers. *Computers in Human Behavior*, *103*, 199-207. https://doi.org/10.1016/j.chb.2019.09.015

Burgess, J., & Green, J. (2009). *YouTube: Online video and participatory culture*. Cambridge Polity Press. https://bit.ly/3Qfi1rs

Burnette, C.B., Kwitowski, M.A., & Mazzeo, S.E. (2017). I don't need people to tell me I'm pretty on social media:" A qualitative study of social media and body image in early adolescent girls. *Body Image*, *23*, 114-125. https://doi.org/10.1016/j.bodyim.2017.09.001

Castillo-Abdul, B., Romero-Rodríguez, L.M., & Larrea-Ayala, A. (2020). Kid influencers in Spain: Understanding the themes they address and preteens' engagement with their YouTube channels. *Heliyon*, *6*(9). https://doi.org/10.1016/j.heliyon.2020.e05056

Davis, K. (2012). Friendship 2.0: Adolescents' experiences of belonging and self-disclosure online. *Journal of Adolescence*, *35*(6), 1527-1536. https://doi.org/10.1016/j.adolescence.2012.02.013

De-Bérail, P., Guillon, M., & Bungener, C. (2019). The relations between YouTube addiction, social anxiety and parasocial relationships with Youtubers: A moderated-mediation model based on a cognitive-behavioral framework. *Computers in Human Behavior*, *99*, 190-204. https://doi.org/10.1016/j.chb.2019.05.007

Du cu, M., & Günneç, D. (2020). Polarity classification of Twitter messages using audio processing. *Information Processing & Management*, *57*(6), 102346-102346. https://doi.org/10.1016/j.ipm.2020.102346

Erz, A., Marder, B., & Osadchaya, E. (2020). Hashtags: Motivational drivers, their use, and differences between influencers and followers. *Computers in Human Behavior*, *89*, 48-60. https://doi.org/10.1016/j.chb.2018.07.030

Ferchaud, A., Grzeslo, J., Orme, S., & Lagroue, J. (2018). Parasocial attributes and YouTube personalities: Exploring content trends across the most subscribed YouTube channels. *Computers in Human Behavior*, *80*, 88-96. https://doi.org/10.1016/j.chb.2017.10.041

Genç, M., & Öksüz, B. (2019). An analysis on collaborations between Turkish beauty YouTubers and cosmetic brands. *Procedia Computer Science*, *158*, 745-750. https://doi.org/10.1016/j.procs.2019.09.110

Harb, J., Ebeling, R., & Becker, K. (2020). A framework to analyze the emotional reactions to mass violent events on Twitter and influential factors. *Information Processing & Management*, *57*(6), 102372. https://doi.org/10.1016/j.ipm.2020.102372

Hartmann, T. (2016). Parasocial interaction, parasocial relationships, and well-being. In L. Reinecke, & M. Oliver (Eds.), *The Routledge handbook of media use and well-being: International perspectives on theory and research on positive media effects* (pp. 131-144). Routledge. https://bit.ly/3zP0GjL

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '04)* (pp. 168-177). Association for Computing Machinery. https://doi.org/10.1145/1014052.1014073

Jerslev, A. (2016). Media times. In the time of the microcelebrity: Celebrification and the YouTuber zoella. *International Journal of Communication*, *10*, 5233-5251. https://bit.ly/3ySPG3r

Kale, G., & Jayanth, J. (2019). Introduction to research. In V. Bairagi, & M. Munot (Eds.), *Research Methodology. A Practical and Scientific Approach*. CRC Press. https://doi.org/10.1201/9781351013277-1

Keegan, B.J., & Rowley, J. (2017). Evaluation and decision making in social media marketing. *Management Decision*, *55*(1), 15-31. https://doi.org/10.1108/MD-10-2015-0450

Kim, D.H., Seely, N.K., & Jung, J.H. (2017). Do you prefer, Pinterest or Instagram? The role of image-sharing SNSs and self-monitoring in enhancing ad effectiveness. *Computers in Human Behavior*, *70*, 535-543. https://doi.org/10.1016/j.chb.2017.01.022

Kim, J., & Kim, Y. (2019). Instagram user characteristics and the color of their photos: Colorfulness, color diversity, and color harmony. *Information Processing & Management*, *56*(4), 1494-1505. https://doi.org/10.1016/j.ipm.2018.10.018

Krippendorff, K. (1980). *Content analysis: An introduction to its methodology*. Sage. https://bit.ly/3bmaPv0

Lange, P.G. (2014). Commenting on YouTube rants: Perceptions of inappropriateness or civic engagement. *Journal of Pragmatics*, *73*, 53-65. https://doi.org/10.1016/j.pragma.2014.07.004

Latorre-Martínez, P., Orive-Serrano, V., & Íñiguez Dieste, D. (2018). Measurement and analysis of the presence in Facebook and Twitter in the regional television broadcaster's context in Spain. *Profesional de la Información*, *27*, 1061-1070. https://doi.org/10.3145/epi.2018.sep.10

León, O.G., & Montero, I. (2015). *Métodos de investigación en Psicología y Educación. Las tradiciones cuantitativa y cualitativa*. McGraw Hill. https://bit.ly/3BCmP6w

Lipsman, A., Mudd, G., Rich, M., & Bruich, S. (2012). The power of "like": How brands reach (and influence) fans through social-media marketing. *Journal of Advertising Research*, *52*(1), 40-52. https://doi.org/10.2501/JAR-52-1-040-052

Lozano-Blasco, R., Quilez-Robres, A., Delgado-Bujedo, D., & Latorre-Martínez, M.P. (2021). YouTube's growth in use among children 0-5 during COVID19: The Occidental European case. *Technology in society*, *66*. https://doi.org/10.1016/j.techsoc.2021.101648

Mäntymäki, M., & Riemer, K. (2014). Digital natives in social virtual worlds: A multi-method study of gratifications and social influences in Habbo Hotel. *International Journal of Information Management*, *34*(2), 210-220. https://doi.org/10.1016/j.ijinfomgt.2013.12.010

Mcgoogan, C. (2017). Hashtag turns 10: Seven facts you didn't know about the trending symbol. *The Telegraph*. https://bit.ly/3coL51b

Neu, D., Saxton, G., Rahaman, A., & Everett, J. (2019). Twitter and social accountability: Reactions to the Panama Papers. *Critical Perspectives on Accounting*, *61*, 38-53. https://doi.org/10.1016/j.cpa.2019.04.003

Nguyen, H., & Le-Nguyen, M. (2018). Multilingual opinion mining on YouTube - A convolutional N-gram BiLSTM word embedding. *Information Processing & Management*, *54*(3), 451-462. https://doi.org/10.1016/j.ipm.2018.02.001

Ofcom (Eds.) (2017). *Children and parents: Media use and attitudes report*. Ofcom. https://bit.ly/3IRiG05

Oramas-Bustillos, R., Zatarain-Cabada, R., Barrón-Estrada, M.L., & Hernández-Pérez, Y. (2019). Opinion mining and emotion recognition in an intelligent learning environment. *Computer Applications in Engineering Education*, *27*(1), 90-101. https://doi.org/10.1002/cae.22059

Peres, R., Talwar, S., Alter, L., Elhanan, M., & Friedmann, Y. (2020). Narrowband influencers and global icons: Universality and media compatibility in the communication patterns of political leaders worldwide. *Journal of International Marketing*, *28*(1), 48-65. https://doi.org/10.1177/1069031X19897893

Reyes-Menéndez, A., Saura, J.R., & Alvarez-Alonso, C. (2018). Understanding #WorldEnvironmentDay user opinions in Twitter: A topic-based sentiment analysis approach. *International Journal of Environmental Research and Public Health*, *15*(11). https://doi.org/10.3390/ijerph15112537

Saura, J.R., Debasa, F., & Reyes-Menendez, A. (2019). Does user generated content characterize Millennials' generation behavior? Discussing the relation between SNS and open innovation. *Journal of Open Innovation: Technology, Market, and Complexity*, *5*(4), 1-15. https://doi.org/10.3390/joitmc5040096

Scannell, P. (2000). For-anyone-as-someone structures. *Media, Culture & Society*, *22*(1), 5-24. https://doi.org/10.1177/016344300022001001

Schmuck, D., Karsay, K., Matthes, J., & Stevic, A. (2019). Looking up and feeling down'. The influence of mobile social networking site use on upward social comparison, self-esteem, and well-being of adult smartphone users. *Telematics and Informatics*, *42*. https://doi.org/10.1016/j.tele.2019.101240

Schouten, A.P., Janssen, L., & Verspaget, M. (2020). Celebrity vs. influencer endorsements in advertising: The role of identification, credibility, and product-endorser fit. *International Journal of Advertising*, *39*(2), 258-281. https://doi.org/10.1080/02650487.2019.1634898

Shane-Simpson, C., Manago, A., Gaggi, N., & Gillespie-Lynch, K. (2018). Why do college students prefer Facebook, Twitter, or Instagram? Site affordances, tensions between privacy and self-expression, and implications for social capital. *Computers in*

Human Behavior, *86*, 276-288. https://doi.org/10.1016/j.chb.2018.04.041

Sharma, S.K., & Hoque, X. (2017). Sentiment predictions using support vector machines for odd-even formula in Delhi. *International Journal of Intelligent Systems and Applications*, *9*(7), 61-69. https://doi.org/10.5815/ijisa.2017.07.07

Smith, A., & Anderson, M. (2018). *Social media use in 2018*. Pew Research Center. https://pewrsr.ch/3aRyOSL

Song, L., Li, R.Y.M., & Yao, Q. (2022). An informal institution comparative study of occupational safety knowledge sharing via French and English Tweets: Languaculture, weak-strong ties and AI sentiment perspectives. *Safety Science*, *147*, 105602. https://doi.org/10.1016/j.ssci.2021.105602

Stockdale, L.A., & Coyne, S.M. (2020). Bored and online: Reasons for using social media, problematic social networking site use, and behavioral outcomes across the transition from adolescence to emerging adulthood. *Journal of Adolescence*, *79*, 173-183. https://doi.org/10.1016/j.adolescence.2020.01.010

Throuvala, M.A., Griffiths, M.D., Rennoldson, M., & Kuss, D.J. (2019). Motivational processes and dysfunctional mechanisms of social media use among adolescents: A qualitative focus group study. *Computers in Human Behavior*, *93*, 164-175. https://doi.org/10.1016/j.chb.2018.12.012

Van-Reijmersdal, E.A., Rozendaal, E., Hudders, L., Vanwesenbeeck, I., Cauberghe, V., & Van-Berlo, Z.M.C. (2020). Effects of disclosing influencer marketing in videos: An eye tracking study among children in early adolescence. *Journal of Interactive Marketing*, *49*(1), 94-106. https://doi.org/10.1016/j.intmar.2019.09.001

Vannucci, A., & Mccauley-Ohannessian, C. (2019). Social media use subgroups differentially predict psychosocial well-being during early adolescence. *Journal of Youth and Adolescence*, *48*, 1469-1493. https://doi.org/10.1007/s10964-019-01060-9

Verrastro, V., Fontanesi, L., Liga, F., Cuzzocrea, F., & Gugliandolo, M.C. (2020). Fear the Instagram: Beauty stereotypes, body image and Instagram use in a sample of male and female adolescents. *Qwerty*, *15*(1), 31-49. https://doi.org/10.30557/QW000021

Vizcaíno-Verdú, A., & Aguaded, I. (2020). Análisis de sentimiento en Instagram: Polaridad y subjetividad de cuentas infantiles. *ZER*, *25*(48), 213-229. https://doi.org/10.1387/zer.21454

Weismueller, J., Harrigan, P., Wang, S., & Soutar, G.N. (2020). Influencer endorsements: How advertising disclosure and source credibility affect consumer purchase intention on social media. *Australasian Marketing Journal*, *28*(4), 160-170. https://doi.org/10.1016/j.ausmj.2020.03.002

Xu, Q.A., Chang, V., & Jayne, C. (2022). A systematic review of social media-based sentiment analysis: Emerging trends and challenges. *Decision Analytics Journal*, *3*. https://doi.org/10.1016/j.dajour.2022.100073

Yau, J.C., & Reich, S.M. (2019). It's just a lot of work": Adolescents' self-presentation norms and practices on Facebook and Instagram. *Journal of Research on Adolescence*, *29*(1), 196-209. https://doi.org/10.1111/jora.12376

Yu, Y., Duan, W., & Cao, Q. (2013). The impact of social and conventional media on firm equity value: A sentiment analysis approach. *Decision Support Systems*, *55*, 919-926. https://doi.org/10.1016/j.dss.2012.12.028

Comunicar

Next Issues

Comunicar 75 (2023-2):
Youth, gender identity and power in digital platforms

Thematic Editors
Dr. Adriana Gewerc, University of Santiago de Compostela (Spain)
Dr. Inés Dusel, Mexican National System of Researchers (SNI) (Mexico)
Dr. David Lee-Carlson, Arizona State University (United States)

Initial call: 01-04-2022
Last call: 30-09-2022

Comunicar 76 (2023-3):
Neurotechnology in the classroom: Current research and future potential

Thematic Editors
Dr. Michael-S.C. Thomas, Birkbeck University of London (United Kingdom)
Dr. Jo-Van Herwegen, UCL Institute of Education (United Kingdom)
Dr. María-José Hernández-Serrano, University of Salamanca (Spain)

Initial call: 01-06-2022
Last call: 30-12-2022

Comunicar 77 (2023-4):
New languages and cultures. Teaching languages for global and digital communication

Thematic Editors
Dr. Anne Pomerantz, University of Pennsylvania (United States)
Dr. Alejandro Rodríguez-Martín, University of Oviedo (Spain)
Dr. Noelia Ibarra-Rius, University of Valencia (Spain)

Initial call: 01-09-2022
Last call: 28-02-2023

Comunicar 78 (2024-1):
Empowered and hyper(dis)connected audiences:
Actors, contexts, experiences and educommunicative practices

Thematic Editors
Dr. José Soengas-Pérez, University of Santiago de Compostela (Spain)
Dr. Amanda Paz-Alencar, Erasmus University Rotterdam (Netherlands)
Dr. Ana-Isabel Rodríguez-Vázquez, University of Santiago de Compostela (Spain)

Initial call: 01-11-2022
Last call: 30-05-2023

Comunicar

9 771113 434785 74

Quarterly scientific journal, bilingual in Spanish and English in all its articles, and abstracts in Portuguese, Chinese and Russian. With 29 years of edition and 1,930 research articles published. Presence in 811 international databases, journal evaluation platforms, selective directories, specialized portals, newspaper catalogues... Rigorous and transparent blind system of evaluation of manuscripts audited in RECYT; International Scientific Board and public network of scientific reviewers of 1090 researchers from 54 countries around the world.

Professional management of manuscripts through the OJS 3 Platform with ethical commitments of transparency and timeliness, anti-plagiarism systems (CrossCheck), alternative metrics (PlumX, Dimensions)... High level of visibility with multiple search systems, DOIs, ORCID, dynamic pdfs, epub and XML, with connection to document managers such as Mendeley, RefWorks, EndNote and scientific social networks such as academia.edu, ResearchGate.

Specialized in educommunication: communication and education, ICT, audiences, new languages...; monographs specialized in current issues. Double format: printed and online; digitally, accessible in full text, free of charge, for the entire scientific community and researchers around the world. Coeditions printed in Spanish and English for the whole world. Published by Comunicar, a non-profit professional association, veteran in Spain (33 years) in education and communication, which collaborates with many international centres and universities.

In active indexations in 2022, Comunicar is top worldwide: 2nd in the world in SCOPUS and 10th in the world in JCR (top 1% and 7% in the world; percentile 99% and 93%). In JCR-JIF it is Q1 in Education, in Communication and in Cultural Studies (1st in Spanish). In SJR is Q1 in Education, Communication and Cultural Studies (1st Spanish in all three areas). It is 1st in FECYT Metrics; 1st in DIALNET METRICS. In GOOGLE SCHOLAR METRICS is the 3rd journal indexed in Spanish in all areas; 2nd in REDIB (out of 1,199 journals).

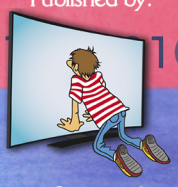
Web of Science Group
JOURNAL CITATION REPORTS
Scopus
FECYT
Google scholar METRICS
DOAJ
DIALNET MÉTRICAS
REDIB
ERIHPLUS
crossref

Sponsored by:

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

novosmedios
grupo de investigación / research group

Published by:



Grupo Comunicar
www.comunicarjournal.com
info@grupocomunicar.com
ISSN: 1134-3478 / e-ISSN: 1988-3293